

Asynchronous CDMA Systems with Random Spreading—Part II: Design Criteria

Laura Cottatellucci, Ralf R. Müller, and Merouane Debbah

Abstract

Totally asynchronous code-division multiple-access (CDMA) systems are addressed. In Part I, the fundamental limits of asynchronous CDMA systems are analyzed in terms of spectral efficiency and SINR at the output of the optimum linear detector. The focus of Part II is the design of low-complexity implementations of linear multiuser detectors in systems with many users. We consider detectors that admit a multistage representation, e.g. reduced rank multistage Wiener filters, polynomial expansion detectors, weighted linear parallel interference cancellers.

The effects of excess bandwidth, chip-pulse shaping, and time delay distribution on CDMA with suboptimum linear receiver structures are investigated. Recursive expressions for universal weight design are given. The performance in terms of SINR is derived in the large-system limit and the performance improvement over synchronous systems is quantified. The considerations distinguish between two ways of forming discrete-time statistics: chip-matched filtering and oversampling.

Index Terms - Asynchronous code division multiple access (CDMA), channel capacity, multiuser detection, random matrix theory, effective interference, linear minimum mean square error (MMSE) detector, multistage detector, random spreading sequences.

This work was presented in part at the IEEE Information Theory Workshop (ITW 2006), Punta de l'Este, Uruguay, March 2006 and at the IEEE Wireless Communications and Networking Conference, Hong Kong, March 2007. It partly appears in Laura Cottatellucci, "Low Complexity Multistage Detectors for Randomly Spread CDMA Systems", Ph.D. thesis, Vienna University of Technology, March 2006.

This work was supported by the French and Australian Science and Technology (FAST) Programme and the Research Council of Norway under grant 171133/V30.

Laura Cottatellucci is with Institute Eurecom, Sophia Antipolis, France (e-mail: laura.cottatellucci@eurecom.fr). She was with Institute of Telecommunications Research, University of South Australia, Adelaide, SA, Australia. Ralf Müller is with Norwegian University of Science and Technology, Trondheim, Norway, (e-mail: ralf.mueller@iet.ntnu.no). Merouane Debbah was with Institut Eurecom, Sophia Antipolis, France. He is currently with SUPELEC, Paris, France (e-mail: merouane.debbah@supelec.fr).

I. INTRODUCTION

In Part I of this paper [1], we analyzed asynchronous CDMA systems with random spreading sequences in terms of spectral efficiency constrained to a given chip pulse waveform and in terms of SINR at the output of an optimum linear multiuser detector. The analysis showed that under realistic conditions, chip-asynchronous CDMA systems significantly outperform chip-synchronous CDMA systems. In order to utilize the benefits from chip-asynchronous¹ CDMA, we need efficient algorithms to cope with multiuser detection for chip-asynchronous users. Therefore, in part II of this work, we focus on the generalization of known design rules for low-complexity multiuser detectors to chip-asynchronous CDMA.

A unified framework for the design and analysis of multiuser detectors that admit a multistage representation for synchronous users was given in [2]. The class of multiuser detectors that admit a multistage representation is large and includes popular linear multiuser detectors like linear MMSE detectors (e.g. [3]), reduced rank multistage Wiener filters [4], [5], polynomial expansion detectors [6] or conjugate gradient methods (e.g. [7]), linear parallel interference cancellers (PIC, e.g. [8], [9]), eventually weighted (e.g. [10]), and the single-user matched filters. Multistage detectors are constructed around the matched filter concept. They consist of a projection of the signal into a subspace of the whole signal space by successive matched filtering and re-spreading followed by a linear filter in the subspace.

Multistage detectors based on universal weights have been proposed in [11], [12] for CDMA systems in AWGN channels and extended to more realistic scenarios in [13], [14], [2]. These references make use of the self-averaging properties of large random matrices to find *universal* weighting coefficients for the linear filter in the subspace. More specifically, the universal weights are obtained by approximating the precise weights designed according to some optimality criterion with asymptotically optimum weights, i.e. the optimum weights for a CDMA system whose number of users and spreading factor tend to infinity with constant ratio. Thanks to the properties of random matrices, asymptotically, these weights become independent of the users' spreading sequences and depend only on few macroscopic system parameters, as the system load or number of transmitted symbols per chip, the variance of the noise, and the distribution of the fading. In this way, the weight design for long-code CDMA simplifies considerably, its complexity becomes independent of both the number of users in the system and the spreading factor. Moreover, the weights need updating only when the macroscopic system parameters change.

¹As already shown in Part I of this paper [1], asynchronism is beneficial when the relative delays between users are *not* integer multiples of a chip interval. To emphasize this requirement we use the term chip-asynchronism instead of asynchronism.

The fact that users are not received in a time-synchronized manner at the receiver causes two main problems from a signal processing perspective: (i) the need for an infinite observation window to implement a linear MMSE detector and (ii) the potential need for oversampling to form sufficient discrete-time statistics. The need for an infinite observation window is primarily related to asynchronism on the symbol-level, not the chip-level. It was addressed in [15], [16] where it was found that multistage detectors need not have infinite observation windows and can be efficiently implemented without windowing at all. A detailed overview of the state of art about statistics, sufficient or not, for multiuser CDMA systems and how to form them was addressed in Part I of this paper [1]. In part I we presented general results with the only constraint that the sampled noise at the output of the front-end was white. For the sake of clarity and to get insights into systems of practical interests, in this part II we focus on two groups of statistics implementable in practical systems:

- (A) Sufficient statistics obtained by filtering the received signal by a lowpass filter with bandwidth B_{LOW} larger than the chip-pulse bandwidth and subsequent sampling at rate $2B_{\text{LOW}}$.
- (B) Statistics obtained by sampling the output of a filter matched to the chip waveform at the chip rate (*chip rate sampling*). In this case, the sampling instants need to be synchronized with the time delay of each user of interest. Thus, different statistics for each user are required. Additionally, the chip pulses at the output of matched filter need to satisfy the Nyquist criterion. In the following we refer to them as root Nyquist chip-pulse waveforms.

General results for the design of linear multistage detectors with both kind of statistics are provided in this work. The chip pulse waveforms are assumed to be identical for all users.

For asynchronous CDMA, low-complexity detectors with universal weights are conveniently obtained from statistics (A). In fact, these observables enable a joint processing of all users without loss of information. Multistage detectors with universal weights and statistics (A) have a complexity order per bit equal to $\mathcal{O}(rK)$ if the sampling rate is $\frac{r}{T_c}$. On the contrary, discretization scheme (B) provides different observables for each user and does not allow for simultaneous joint detection of all users. An implementation of multistage detectors with universal weights using such statistics implies a complexity order *per bit* equal to $\mathcal{O}(K^2)$. This approach is still interesting from a complexity point of view if detection of a single user is required. However, it suffers from a performance degradation due to the sub-optimality of the statistics.

This work is organized in six additional sections. Section II and III introduce the notation and the system model for asynchronous CDMA, respectively. In Section IV, multistage detectors for asynchronous CDMA

are reviewed and a implementation which does not suffer from truncation effects is given. The design of universal weighting is addressed in Section V. Finally, the analytical results are applied to gain further insight into the system in Section VI where methods for pulse-shaping, forming sufficient statistics and synchronization are compared. Conclusions are summed up in Section VII.

II. NOTATION AND SOME USEFUL DEFINITIONS

Throughout Part II we adopt the same notation and definitions already introduced in Part I of this work [1]. In order to make Part II self-contained we repeat here definitions useful in this part. Upper and lower boldface symbols are used respectively for matrices and vectors corresponding to signals spanning a specific symbol interval m . Matrices and vectors describing signals spanning more than a symbol interval are denoted by upper boldface calligraphic letters.

In the following, we utilize *unitary* Fourier transforms both in the continuous time and in the discrete time domain. The unitary Fourier transform of a function $f(t)$ in the continuous time domain is given by $F(\omega) = \frac{1}{\sqrt{2\pi}} \int f(t) e^{-j\omega t} dt$. The unitary Fourier transform of a sequence $\{\dots, c_{-1}, c_0, c_1, \dots\}$ in the discrete time domain is given by $c(\Omega) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{+\infty} c_n e^{-j\Omega n}$. We will refer to them shortly as Fourier transform. We denote the argument of a Fourier transform of a continuous function by ω and the argument of a Fourier transform of a sequence by Ω . They are the angular frequency and the normalized angular frequency, respectively. A function in Ω is periodic with respect to integer multiples of 2π .

For further studies it is convenient to define the concept of *r-block-wise circulant matrices of order N*.

Definition 1 Let r and N be positive integers. An *r-block-wise circulant matrix of order N* is an $rN \times N$ matrix of the form

$$\mathbf{C} = \begin{pmatrix} \mathbf{B}_0 & \mathbf{B}_1 & \cdots & \mathbf{B}_{N-1} \\ \mathbf{B}_{N-1} & \mathbf{B}_0 & \cdots & \mathbf{B}_{N-2} \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_0 \end{pmatrix} \quad (1)$$

with $\mathbf{B}_i = (c_{1,i}, c_{2,i}, \dots, c_{r,i})^T$.

In the matrix \mathbf{C} an $r \times N$ block row is obtained by circularly right shift of the previous block. Since the matrix \mathbf{C} is univocally defined by the unitary Fourier transforms of the sequences $\{c_{s,0}, c_{s,1}, \dots, c_{s,N-1}\}$, for

$s = 1 \dots r$,

$$c_s(\Omega) = \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{N-1} c_{sk} e^{-j\Omega k} \quad s = 1, \dots, r,$$

there exists a bijection \mathfrak{F} from the frequency dependent vector $\mathbf{c}(\Omega) = [c_1(\Omega), c_2(\Omega), \dots, c_r(\Omega)]$ to \mathbf{C} . Thus,

$$\mathbf{C} = \mathfrak{F}\{\mathbf{c}(\Omega)\}. \quad (2)$$

Furthermore, the superscripts \cdot^T , \cdot^H , and \cdot^* , denote the transpose, the conjugate transpose, and the conjugate of the matrix argument, respectively. \mathbf{I}_n is the identity matrix of size $n \times n$ and \mathbb{C} , \mathbb{Z} , \mathbb{Z}^+ , \mathbb{N} , and \mathbb{R} are the fields of complex, integer, nonnegative integers, natural, and real numbers, respectively. $\text{tr}(\cdot)$ is the trace of the matrix argument and $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_s)$ denotes the vector space spanned by the s vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_s$. $\text{diag}(\dots) : \mathbb{C}^n \rightarrow \mathbb{C}^{n \times n}$ transforms an n -dimensional vector \mathbf{v} into a diagonal matrix of size n having as diagonal elements the components of \mathbf{v} in the same order. $\mathbb{E}\{\cdot\}$ and $\text{Pr}\{\cdot\}$ are the expectation and probability operators, respectively. δ_{ij} is the Kronecker symbol and $\delta(\lambda)$ is the Dirac's delta function. mod denotes the modulus and $\lfloor \cdot \rfloor$ is the operator that yields the maximum integer not greater than its argument.

III. SYSTEM MODEL

In this section we recall briefly the system model for asynchronous CDMA introduced in Section IV and VII of Part I of this work [1]. The reader interested in the details of the derivation can refer to [1].

Let us consider an asynchronous CDMA system with K active users in the uplink channel with spreading factor N . Each user and the base station are equipped with a single antenna. The channel is flat fading and impaired by additive white Gaussian noise with power spectral density N_0 . The symbol interval is denoted with T_s and $T_c = \frac{T_s}{N}$ is the chip interval. The modulation of all users is based on the same chip pulse waveform $\psi(t)$ bandlimited with bandwidth B , unitary Fourier transform $\Psi(\omega)$, and energy $E_\psi = \int_{-\infty}^{\infty} |\psi(t)|^2 dt$.

The time delays of the K users are denoted with τ_k , $k = 1, \dots, K$. Without loss of generality we can assume (i) user 1 as reference user so that $\tau_1 = 0$, (ii) the users ordered according to increasing time delay with respect to the reference user, i.e. $\tau_1 \leq \tau_2 \leq \dots \leq \tau_K$; (iii) the time delay to be, at most, one symbol interval so that $\tau_k \in [0, T_s)$.²

As for the results presented in Part I, the mathematical results presented in this second part hold for any front-end that keeps the sampled noise white at its output. However, in order to get better insights into

²For a thorough discussion on this assumption the reader can refer to [3].

the physical system we focus on two front-ends of practical and theoretical interest. Both of them satisfy the more general assumption underlying the results in Part I. We refer to them as Front-end Type A and Front-end Type B³.

Front-end Type A consists of

- An ideal lowpass filter with cut-off frequency $\omega = \frac{\pi r}{T_c}$ where $r \in \mathbb{Z}^+$ satisfies the constraint $B \leq \frac{r}{2T_c}$ such that the sampling theorem applies. The filter is normalized to obtain a unit overall amplification factor, i.e. the transfer function is

$$G(\omega) = \begin{cases} \frac{1}{\sqrt{E_\psi}} & |\omega| \leq \frac{\pi r}{T_c} \\ 0 & |\omega| > \frac{\pi r}{T_c}. \end{cases} \quad (3)$$

- A subsequent continuous-discrete time conversion by sampling at rate $\frac{r}{T_c}$.

This front-end satisfies the conditions of the sampling theorem and, thus, provides sufficient discrete-time statistics. For convenience, the sampling rate is an integer multiple of the chip rate. Additionally, the discrete-time noise process is white with zero mean and variance $\sigma^2 = \frac{N_0 r}{E_\psi T_c}$.

Front-end Type B consists of

- A filter $G(\omega)$ matched to the chip pulse and normalized to the chip pulse energy, i.e. $G(\omega) = \Psi^*(\omega)E_\psi^{-\frac{1}{2}}$;
- Subsequent sampling at the chip rate.

When used with root Nyquist chip pulses, the discrete time noise process $\{w[p]\}$ is white with variance $\frac{N_0}{E_\psi T_c}$. For a synchronous systems with square root Nyquist chip pulses, this front end provides sufficient statistics whereas the observables are not sufficient if the system is asynchronous.

The chip waveform at the filter output is denoted by $\phi(t)$ and its unitary Fourier transform by $\Phi(\omega)$. The well-known relations $\phi(t) = \psi(t) * g(t)$ and $\Phi(\omega) = \Psi(\omega)G(\omega)$ hold. The unitary Fourier transform of the chip pulse waveform $\phi(t)$ sampled at rate $\frac{1}{T_c}$ and delay τ is given by

$$\phi(\Omega, \tau) \triangleq \frac{1}{T_c} \sum_{s=-\infty}^{+\infty} e^{j\frac{\tau}{T_c}(\Omega+2\pi s)} \Phi^* \left(\frac{j(\Omega+2\pi s)}{T_c} \right). \quad (4)$$

Sufficient statistics for asynchronous CDMA require an infinite observation window. In the following, we introduce a matrix system model corresponding to an infinite observation window.

³For the sake of compactness of some of the results, we adopt another normalization than in Part I. Here, the signal energy at the output of the front-end is equal to one. In Part I, the energy of the analog filter's impulse response is normalized to unity. The variance of the sampled noise at the front-end output changes accordingly.

Let us denote with $\mathbf{b}^{(m)}$ and $\mathbf{y}^{(m)}$ the vectors of transmitted and received signals at time instants $m \in \mathbb{Z}$. The baseband discrete-time asynchronous system is given by

$$\mathcal{Y} = \mathcal{H}\mathcal{B} + \mathcal{W} \quad (5)$$

where $\mathcal{Y} = [\dots, \mathbf{y}^{(m-1)T}, \mathbf{y}^{(m)T}, \mathbf{y}^{(m+1)T} \dots]^T$ and $\mathcal{B} = [\dots, \mathbf{b}^{(m-1)T}, \mathbf{b}^{(m)T}, \mathbf{b}^{(m+1)T} \dots]^T$ are infinite-dimensional vectors of received and transmitted symbols respectively; \mathcal{W} is an infinite-dimensional noise vector; and \mathcal{H} is a bi-diagonal block matrix of infinite size given by

$$\mathcal{H} = \begin{bmatrix} \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots \\ \dots & \mathbf{0} & \mathbf{H}_d^{(m-1)} & \mathbf{H}_u^{(m)} & \mathbf{0} & \dots & \dots \\ \dots & \dots & \mathbf{0} & \mathbf{H}_d^{(m)} & \mathbf{H}_u^{(m+1)} & \mathbf{0} & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \quad (6)$$

Here, $\mathbf{H}_u^{(m)}$ and $\mathbf{H}_d^{(m)}$ are matrices of size $rN \times K$ obtained by the decomposition of the $2rN \times K$ matrix $\mathbf{H}^{(m)}$ into two parts such that $\mathbf{H}^{(m)} = [\mathbf{H}_u^{(m)T}, \mathbf{H}_d^{(m)T}]^T$. For $\mathbf{H}^{(m)}$ the relation

$$\mathbf{H}^{(m)} = \mathbf{S}^{(m)} \mathbf{A} \quad (7)$$

holds where \mathbf{A} is the $K \times K$ diagonal matrix of the received amplitudes a_k and $\mathbf{S}^{(m)}$ is the $2rN \times K$ matrix whose k -th column accounts for the spreading of the symbol transmitted by user k in the symbol interval m and due to the actual spreading sequence, the channel delay, and filtering and sampling at the front-end. We refer to it as the matrix of virtual spreading. More specifically, the matrix of virtual spreading is given by

$$\mathbf{S}^{(m)} = \left(\Phi_1 \mathbf{s}_1^{(m)}, \Phi_2 \mathbf{s}_2^{(m)}, \dots, \Phi_K \mathbf{s}_K^{(m)} \right) \quad (8)$$

where $\mathbf{s}_k^{(m)}$ is the N -dimensional column vector of the spreading sequence of user k for the transmitted symbol m and Φ_k is the $2rN \times N$ matrix taking into account the effects of the chip pulse shape and the time delay τ_k user k . Let us decompose τ_k in $\bar{\tau}_k = \left\lfloor \frac{\tau_k}{T_c} \right\rfloor$ and $\tilde{\tau}_k = \tau_k - T_c \bar{\tau}_k = \tau_k \bmod T_c$, the integer number of chips the signal is delayed and its delay within a chip, respectively. The matrix Φ_k is of the form

$$\Phi_k = \begin{bmatrix} \mathbf{0}_{\bar{\tau}_k} \\ \tilde{\Phi}_k \\ \mathbf{0}_{N-\bar{\tau}_k} \end{bmatrix} \quad (9)$$

where $\mathbf{0}_{\bar{\tau}_k}$ and $\mathbf{0}_{N-\bar{\tau}_k}$ are zero matrices of dimensions $\bar{\tau}_k \times N$ and $(N - \bar{\tau}_k) \times N$, respectively; $\tilde{\Phi}_k$ is an r -block-wise circulant matrix of order N as in (2)

$$\tilde{\Phi}_k = \mathfrak{F}(\mathbf{c}(\tilde{\tau}_k)), \quad (10)$$

with

$$\mathbf{c}(\tilde{\tau}_k) = \left[\phi(\Omega, \tilde{\tau}_k) \phi(\Omega, \tilde{\tau}_k - \frac{T_c}{r}), \dots, \phi(\Omega, \tilde{\tau}_k - \frac{(r-1)T_c}{r}) \right].$$

Thus, the virtual spreading sequences are the samples of the delayed continuous-time spreading waveforms at sampling rate r/T_c .

Throughout this work we assume that the transmitted symbols are uncorrelated and identically distributed random variables with unitary variance and zero mean, i.e. $\mathbb{E}(\mathbf{B}) = \mathbf{0}$ and $\mathbb{E}(\mathbf{B}\mathbf{B}^H) = \mathbf{I}$ being $\mathbf{0}$ and \mathbf{I} the unlimited zero vector and the unlimited identity matrix, respectively. The elements of the spreading sequences $\mathbf{s}_k^{(m)}$ are assumed to be zero mean i.i.d. Gaussian random variables over all the users, chips, and symbols with $\mathbb{E}\{\mathbf{s}_k^{(m)} \mathbf{s}_k^{(m)H}\} = \frac{1}{N} \mathbf{I}_N$. Finally, $\mathbf{u}_k^{(m)}$ denotes that column of the matrix \mathcal{H} containing the k^{th} column of the matrix $\mathbf{H}^{(m)}$. We define the correlation matrices $\mathcal{T} = \mathcal{H}\mathcal{H}^H$ and $\mathcal{R} = \mathcal{H}^H\mathcal{H}$. The system load $\beta = \frac{K}{N}$ is the number of transmitted symbols per chip.

IV. MULTISTAGE STRUCTURES FOR ASYNCHRONOUS CDMA

We consider the large class of linear multistage detectors for asynchronous CDMA. Let $\chi_{L,k}^{(m)}(\mathcal{H})$ be the Krylov subspace [17] of rank $L \in \mathbb{Z}^+$ given by

$$\chi_{L,k}^{(m)}(\mathcal{H}) = \text{span}(\mathcal{T}^\ell \mathbf{u}_k^{(m)})_{\ell=0}^{L-1}. \quad (11)$$

A multistage detector of rank $L \in \mathbb{Z}^+$ for user k is given by

$$\hat{b}_k = \sum_{\ell=0}^{L-1} (\mathbf{w}_k^{(m)})_\ell \mathbf{u}_k^{(m)H} \mathcal{T}^\ell \mathbf{y} \quad (12)$$

where $\mathbf{w}_k^{(m)}$ is the L -dimensional vector of weight coefficients.

It has been shown in [16] that, given the weight vector $\mathbf{w}_k^{(m)}$ the detection of the symbol $b_k^{(m)}$ by the multistage detector of rank L in (12) can be performed with finite delay L using the implementation scheme in Figure 1. Although infinite length vectors and infinite dimension matrices appear in (12), the multistage detector in Figure 1 implements exactly (12) and does not suffer from truncation effects. Equivalently, the multistage detector in Figure 1 can be considered as a multistage detector processing data over an observation window of size $2L$. The projection of the received vector \mathbf{y} onto the subspaces $\chi_{L,k}^{(m)}(\mathcal{H})$, for $k = 1 \dots K$, is performed jointly for all users and requires only multiplications between vectors and matrices. The size of those vectors and matrices does not depend on the observation window. For further details the interested reader is referred to [16], [18].

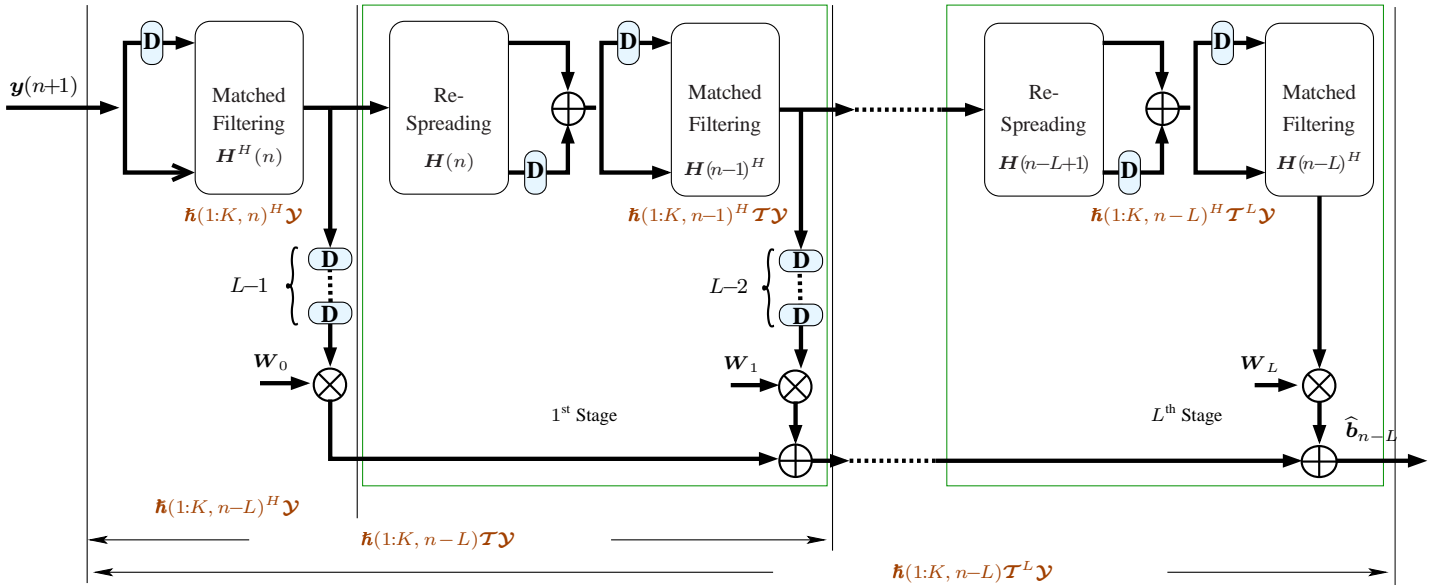


Fig. 1. Multistage detector for asynchronous CDMA systems. Here, $\mathbf{h}(1:K, n) = [\Phi_1 s_1^{(n)}, \Phi_2 s_2^{(n)}, \dots, \Phi_K s_K^{(n)}]$

The class of multistage detectors includes many popular multiuser detectors:

- the single-user matched filter for $L = 1$,
- the linear parallel interference canceller (PIC) [19], [20] for weight coefficients chosen irrespective of the properties of the transfer matrix \mathcal{H} ,
- the polynomial expansion detector [6] and the conjugate gradient method [7], if the weight coefficients are identical for all users and chosen to minimize the mean square error,
- the (reduced rank) multistage Wiener filter [5] if the weight coefficients are chosen to minimize the mean square error, but are allowed to differ from user to user.

Throughout this work we refer to detectors that minimize the MSE in the projection subspace of the user of interest as *optimum detectors in the MSE sense*. More specifically this class of multistage detectors includes the linear MMSE detector and the multistage Wiener filter but not the polynomial expansion detector.

In the following we focus on the design of multistage Wiener filters implemented as in Figure 1. This reduces the problem to the design of the filter coefficients $\mathbf{w}_k^{(m)}$. The multistage Wiener filter for the detection of the symbol m transmitted by user k reads

$$\mathcal{M}_k^{(m)} = \sum_{\ell=0}^{L-1} (\mathbf{w}_k^{(m)})_{\ell-1} \mathbf{u}_k^{(m)H} \mathbf{T}^\ell. \quad (13)$$

The weight vector $\mathbf{w}_k^{(m)}$ that minimizes the MSE $E\{\|\mathcal{M}_k^{(m)}\mathbf{y} - b_k^{(m)}\|^2\}$ is given by

$$\mathbf{w}_k^{(m)} = \underset{\bar{\mathbf{w}}_k^{(m)}}{\operatorname{argmin}} E \left\{ \left\| \sum_{\ell=0}^{L-1} (\bar{\mathbf{w}}_k^{(m)})_{\ell} \mathbf{u}_k^{(m)H} \mathcal{T}^{\ell} \mathbf{y} - b_k^{(m)} \right\|^2 \right\} \quad (14)$$

$$= \underset{\bar{\mathbf{w}}_k^{(m)}}{\operatorname{argmin}} E \left\{ \left\| \bar{\mathbf{w}}_k^{(m)H} \mathbf{x}_k^{(m)} - b_k^{(m)} \right\|^2 \right\} \quad (15)$$

where $\mathbf{x}_k^{(m)}$ is an L -dimensional vector with j^{th} element $(\mathbf{x}_k^{(m)})_j = \mathbf{u}_k^{(m)H} \mathcal{T}^{j-1} \mathbf{y}$. This optimization problem is solved by the Wiener-Hopf theorem [21] and $\mathbf{w}_k^{(m)}$ is given by

$$\mathbf{w}_k^{(m)} = (\Xi_k^{(m)})^{-1} \boldsymbol{\xi}_k^{(m)} \quad (16)$$

where $\Xi_k^{(m)} = E\{\mathbf{x}_k^{(m)} \mathbf{x}_k^{(m)H}\}$ and $\boldsymbol{\xi}_k^{(m)} = E\{b_k^{(m)*} \mathbf{x}_k^{(m)}\}$. It is straightforward to verify that in this case

$$\Xi_k^{(m)} = \begin{pmatrix} (\mathcal{R}^2)_{k,m} + \sigma^2(\mathcal{R})_{k,m} & \cdots & (\mathcal{R}^{L+1})_{k,m} + \sigma^2(\mathcal{R}^L)_{k,m} \\ (\mathcal{R}^3)_{k,m} + \sigma^2(\mathcal{R}^2)_{k,m} & \cdots & (\mathcal{R}^{L+2})_{k,m} + \sigma^2(\mathcal{R}^{L+1})_{k,m} \\ \vdots & \ddots & \vdots \\ (\mathcal{R}^{L+1})_{k,m} + \sigma^2(\mathcal{R}^L)_{k,m} & \cdots & (\mathcal{R}^{2L})_{k,m} + \sigma^2(\mathcal{R}^{2L-1})_{k,m} \end{pmatrix}$$

$$\boldsymbol{\xi}_k^{(m)} = ((\mathcal{R})_{k,m}, (\mathcal{R}^2)_{k,m}, \dots, (\mathcal{R}^L)_{k,m})^T. \quad (17)$$

where $(\mathcal{R}^s)_{k,m} = \mathbf{h}_k^{(m)H} \mathcal{T}^{s-1} \mathbf{h}_k^{(m)}$ is the diagonal element of the matrix \mathcal{R}^s corresponding to the m^{th} symbol transmitted by user k .

V. UNIVERSAL WEIGHT DESIGN

Consider the SINR of any linear detector that admits a multistage representation. Let $\bar{\mathbf{w}}_{k,m}$ be the weight vector for the detection of the m^{th} symbol transmitted by user k . Then the SINR at the output of the multistage detector is given by

$$\text{SINR}_k = \frac{\bar{\mathbf{w}}_k^{(m)H} \boldsymbol{\xi}_k^{(m)} \boldsymbol{\xi}_k^{(m)T} \bar{\mathbf{w}}_k^{(m)}}{\bar{\mathbf{w}}_k^{(m)H} (\Xi_k^{(m)} - \boldsymbol{\xi}_k^{(m)} \boldsymbol{\xi}_k^{(m)T}) \bar{\mathbf{w}}_k^{(m)H}}. \quad (18)$$

The performance of multistage Wiener filters simplifies to

$$\text{SINR}_k = \frac{\boldsymbol{\xi}_k^{(m)T} \Xi_k^{(m)-1} \boldsymbol{\xi}_k^{(m)}}{1 - \boldsymbol{\xi}_k^{(m)T} \Xi_k^{(m)-1} \boldsymbol{\xi}_k^{(m)}}. \quad (19)$$

From (16), (18), and (19) it is apparent that the diagonal elements of the matrix \mathcal{R}^s play a fundamental role in the design and analysis of multistage detectors.

It has been shown in [2] that, if the spreading sequences are random and the CDMA system is synchronous, the diagonal elements of the matrix \mathcal{R}^s , $s \in \mathbb{Z}^+$, converge to deterministic values as $K, N \rightarrow \infty$ with constant ratio. This asymptotic convergence holds for some classes of random matrices and is a stronger property than the convergence of the eigenvalue distribution. The Stieltjes transform of the asymptotic eigenvalue distribution of \mathcal{R} is related to the SINR at the output of the linear MMSE detector, as pointed out first in [22] for synchronous CDMA systems. The asymptotic eigenvalue moments of \mathcal{R} enable the asymptotic performance analysis of reduced rank multistage Wiener filters [23] and the design of multistage detectors with quadratic complexity order per bit [14], [13]. The convergence of the diagonal elements of \mathcal{R}^s has been utilized in [2] for the design of multistage detectors with linear complexity order per bit in synchronous CDMA systems and for the asymptotic analysis of any multistage detector not necessarily optimum in a MSE sense. In the following we extend the results in [2] to the case of asynchronous CDMA systems making use of the asymptotic properties of the random matrix \mathcal{R} for asynchronous CDMA systems.

The design of low complexity multistage detectors is based on the approximation of the weight vectors $\mathbf{w}_k^{(m)}$ by their asymptotic limit when $K, N \rightarrow \infty$ with constant ratio β

$$\mathbf{w}_k^\infty = \lim_{K=\beta N \rightarrow \infty} \mathbf{\Xi}_k^{(m)-1} \mathbf{\xi}_k^{(m)}. \quad (20)$$

Thanks to the fact that the diagonal elements of \mathcal{R}^s can be computed by a polynomial in few macroscopic system parameters, the computation of the weight vectors becomes independent of the size of \mathcal{R} and independent of m . Thus, the effort for the computation of the weights becomes negligible and the complexity of the detector is dominated by the joint projection of the received signal \mathcal{Y} onto the subspaces $\chi_k^{(m)}(\mathcal{H})$, $k = 1 \dots K$ and $m \in \mathbb{Z}$. This projection has linear complexity per bit if the multistage detector in Figure 1 is utilized.

The convergence of the diagonal elements of \mathcal{R}^ℓ to deterministic values is established in the following theorem. The definitions and the assumptions in the statement of Theorem 1 summarize and formalize the characteristics of system model (5) for $\tau_k \in [0, T_s]$.

Theorem 1 *Let $K, N \in \mathbb{N}$ and $\mathbf{A} \in \mathbb{C}^{K \times K}$ be a diagonal matrix with k^{th} diagonal element $a_k \in \mathbb{C}$. T_s and T_c are positive reals with $T_s = NT_c$. Given $\{\tau_1, \tau_2, \dots, \tau_K\}$ a set of delays in $[0, T_s)$, we introduce the sets of delays in $[0, T_c)$ defined as $\{\tilde{\tau}_k : \tilde{\tau}_k = \tau_k \bmod T_c, k = 1, \dots, K\}$ and the set of normalized delays $\{\bar{\tau}_k : \bar{\tau}_k = \left\lfloor \frac{\tau_k}{T_c} \right\rfloor\}$. Given a function $\Phi(\omega) : \mathbb{R} \rightarrow \mathbb{C}$, let $\phi(\Omega, \tau)$ be as in (4). Given*

a positive integer r , let Φ_k , $k = 1, \dots, K$, be r -block-wise circulant matrices of order N defined in (10) and $\mathbf{S}^{(m)} = (\Phi_1 \mathbf{s}_1^{(m)}, \Phi_2 \mathbf{s}_2^{(m)}, \dots, \Phi_K \mathbf{s}_K^{(m)})$ with $\mathbf{s}_k^{(m)}$ N -dimensional random column vector. Let $\mathbf{H} = (\mathbf{H}_u^{(m)T}, \mathbf{H}_d^{(m)T})^T = \mathbf{S}\mathbf{A}$ with $\mathbf{H}_u^{(m)}, \mathbf{H}_d^{(m)} \in \mathbb{C}^{rN \times K}$ and \mathcal{H} the infinite block row and block column matrix of the same form as in (6), $\mathcal{T} = \mathcal{H}\mathcal{H}^H$, $\mathcal{R} = \mathcal{H}^H\mathcal{H}$, and $\mathbf{u}_k^{(m)}$ the column of \mathcal{H} corresponding to $\Phi_k \mathbf{s}_k^{(m)}$.

We assume that the function $\Phi(\omega)$ is upper bounded and has finite support. The receive filter is such that the sampled discrete time noise process is white. The vectors \mathbf{s}_k are independent with i.i.d. zero-mean circularly symmetric Gaussian elements with variance $\mathbb{E}\{|s_{ij}|^2\} = N^{-1}$. Furthermore, the elements a_k of the matrix \mathbf{A} are uniformly bounded for any K . The sequence of the empirical joint distributions $F_{|\mathbf{A}|^2, \tilde{\tau}}^{(K)}(\lambda, \tilde{\tau}) = \frac{1}{K} \sum_{k=1}^K 1(\lambda - |a_k|^2) 1(\tilde{\tau} - \tilde{\tau}_k)$ converges almost surely, as $K \rightarrow \infty$, to a non-random distribution function $F_{|\mathbf{A}|^2, \tilde{\tau}}(\lambda, \tilde{\tau})$.

Then, conditioned on $(|a_k|^2, \tilde{\tau}_k)$, the corresponding diagonal elements of the matrices \mathcal{R}^ℓ converge almost surely to the deterministic value

$$\lim_{K=\beta N \rightarrow \infty} (\mathcal{R}^\ell)_{k,m} = \lim_{K=\beta N \rightarrow \infty} \mathbf{u}_k^{(m)H} \mathcal{T} \mathbf{u}_k^{(m)} \stackrel{a.s.}{=} R_\ell(|a_k|^2, \tilde{\tau}_k) \quad (21)$$

with $R_\ell(|a_k|^2, \tilde{\tau}_k)$ determined by the following recursion

$$R_\ell(\lambda, \tau) = \sum_{s=0}^{\ell-1} g(\mathbf{T}_{\ell-s-1}, \lambda, \tau) R_s(\lambda, \tau) \quad (22)$$

and

$$\mathbf{T}_\ell(\Omega) = \sum_{s=0}^{\ell-1} \mathbf{f}(R_{\ell-s-1}, \Omega) \mathbf{T}_s(\Omega) \quad -\pi \leq \Omega \leq \pi \quad (23)$$

$$\mathbf{f}(R_\ell, \Omega) = \beta \int \lambda \Delta_{\phi,r}(\Omega, \tau) \Delta_{\phi,r}^H(\Omega, \tau) R_\ell(\lambda, \tau) dF_{|\mathbf{A}|^2, T}(\lambda, \tau) \quad -\pi \leq \Omega \leq \pi \quad (24)$$

$$g(\mathbf{T}_\ell, \lambda, \tau) = \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \Delta_{\phi,r}^H(\Omega, \tau) \mathbf{T}_\ell(\Omega) \Delta_{\phi,r}(\Omega, \tau) d\Omega \quad (25)$$

with

$$\Delta_{\phi,r}(\Omega, \tau) = \begin{pmatrix} \phi(\Omega, \tau) \\ \phi(\Omega, \tau - \frac{T_c}{r}) \\ \vdots \\ \phi(\Omega, \tau - \frac{T_c(r-1)}{r}) \end{pmatrix}. \quad (26)$$

The recursion is initialized by setting $\mathbf{T}_0(\Omega) = \mathbf{I}_r$ and $R_0(\lambda, \tau) = 1$.

Theorem 1 is proven in Appendix I.

Note that the asymptotic diagonal elements of \mathcal{R}^ℓ depend on the delay τ_k only via the delay of a chip pulse waveform within a chip, i.e. via $\tilde{\tau}_k$, while any delay multiple of T_c leaves the diagonal elements unchanged.

From Theorem 1 we can obtain $m_{\mathcal{R}}^{(\ell)}$, the asymptotic eigenvalue moment of the matrix \mathcal{R} of order ℓ by using the relation

$$m_{\mathcal{R}}^{(\ell)} = \mathbb{E}\{R_\ell(\lambda, \tau)\}$$

where the expectation is taken over the limit distribution $F_{|\mathbf{A}|^2, \tilde{T}}(\lambda, \tilde{\tau})$. For $r = 1$ and $F_{|\mathbf{A}|^2, \tilde{T}}(\lambda, \tilde{\tau}) = F_{|\mathbf{A}|^2}(\lambda)\delta(\tilde{\tau})$, i.e. for synchronous systems sampled at the chip rate, and $\Phi(\omega)$ satisfying the Nyquist criterion the recursive equations (23), (24), and (25) reduce to the recursion in [2] Theorem 1.

This theorem is very general and holds for all chip pulses of practical interest. Furthermore, no constraint is imposed on the time delay distribution. The choice of the front end in this work is restricted only by the applicability of (18) or (19), which imply white noise at the front end. Then, since both Front-end A and Front -end B keep the sampled noise white, Theorem 1 applies to both of them.

Now, we specialize Theorem 1 to a case of theoretical and practical interest, where sufficient statistics are utilized in the detection, the chip pulse waveform $\phi(t)$ is band-limited, and the sequence of the empirical distribution functions of the time delays converges to a uniform distribution function as $K \rightarrow +\infty$. The constraint to use sufficient statistics restricts the class of front-ends. The following results apply to Front-end A but, in general, not to Front-end B.

Corollary 1 *Let us adopt the same definitions as in Theorem 1 and let the same assumptions of Theorem 1 be satisfied. Additionally, assume that the random variables λ and $\tilde{\tau}$ in $F_{|\mathbf{A}|^2, \tilde{T}}(\lambda, \tilde{\tau})$ are statistically independent and the random variable $\tilde{\tau}$ is uniformly distributed. Furthermore, $\Phi(\Omega)$ is bounded in absolute*

value, and bandlimited with bandwidth $B \leq \frac{r}{2T_c}$. Then, given $(|a_k|^2, \tilde{\tau}_k)$ and $m \in \mathbb{Z}$, the corresponding diagonal element of the matrix \mathcal{R}^ℓ converges almost surely to a deterministic value, conditionally on $|a_k|^2$,

$$\lim_{K=\beta N \rightarrow \infty} (\mathcal{R}^\ell)_{k,m} = \lim_{K=\beta N \rightarrow \infty} \mathbf{u}_k^{(m)H} \mathbf{T}^{\ell-1} \mathbf{u}_k^{(m)} \stackrel{a.s.}{=} R_\ell(|a_k|^2)$$

with $R_\ell(\lambda)|_{\lambda=|a_k|^2}$ determined by the following recursion:

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} \lambda R_s(\lambda) \nu_{\ell-s-1}$$

and

$$\begin{aligned} T_\ell(\omega) &= \frac{r}{T_c} \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \frac{1}{T_c} |\Phi(\omega)|^2 T_s(\omega) & -2\pi B \leq \omega \leq 2\pi B \\ f(R_\ell) &= \beta \int \lambda R_\ell(\lambda) dF_{|\mathbf{A}|^2}(\lambda) \\ \nu_\ell &= \frac{r}{2\pi T_c} \int_{-2\pi B}^{2\pi B} |\Phi(\omega)|^2 T_\ell(\omega) d\omega. \end{aligned}$$

The recursion is initialized by setting $T_0(\omega) = 1$ and $R_0(\lambda) = 1$.

Corollary 1 is derived in Appendix II.

The eigenvalue moments of \mathcal{R} can be expressed in terms of the auxiliary quantities $f(R_s)$ and ν_s in the recursion of Corollary 1 by the following expression:

$$m_{\mathcal{R}}^{(\ell)} = \mathbb{E}\{R_\ell(\lambda)\} = \sum_{s=0}^{\ell-1} f(R_s) \nu_{\ell-s-1}.$$

Applying Corollary 1 we obtain the following algorithm to compute the asymptotic limits of the diagonal elements of \mathcal{R}^ℓ and its eigenvalue moments.

Algorithm 1

Initialization: Let $\rho_0(z) = 1$ and $\mu_0(y) = 1$.

- l^{th} step:
- Define $u_{\ell-1}(y) = ry\mu_{\ell-1}(y)$ and write it as a polynomial in y .
 - Define $v_{\ell-1}(z) = z\rho_{\ell-1}(z)$ and write it as a polynomial in z .

- *Define*

$$\mathcal{E}_s = \frac{1}{2\pi T_c^s} \int_{-2\pi B}^{2\pi B} T_c |\Phi(\omega)|^{2s} d\omega \quad (27)$$

and replace all monomials y, y^2, \dots, y^ℓ in the polynomial $u_{\ell-1}(y)$ by $\mathcal{E}_1/T_c, \mathcal{E}_2/T_c, \dots, \mathcal{E}_\ell/T_c$, respectively. Denote the result by $U_{\ell-1}$.

- *Define $m_{|\mathbf{A}|^2}^s = \mathbb{E}\{|a_k|^{2s}\}$ and replace all monomials z, z^2, \dots, z^ℓ in the polynomial $v_{\ell-1}(z)$ by the moments $m_{|\mathbf{A}|^2}^{(1)}, m_{|\mathbf{A}|^2}^{(2)}, \dots, m_{|\mathbf{A}|^2}^{(\ell)}$, respectively. Denote the result by $V_{\ell-1}$.*
- *Calculate*

$$\begin{aligned} \rho_\ell(z) &= \sum_{s=0}^{\ell-1} z U_{\ell-s-1} \rho_s(z) \\ \mu_\ell(y) &= \frac{r}{T_c} \sum_{s=0}^{\ell-1} \beta y V_{\ell-s-1} \mu_s(y). \end{aligned}$$

- *Assign $\rho_\ell(\lambda)$ to $R^\ell(\lambda)$.*

Replace all monomials z, z^2, \dots, z^ℓ in the polynomial $\rho_\ell(z)$ by the moments $m_{|\mathbf{A}|^2}^{(1)}, m_{|\mathbf{A}|^2}^{(2)}, \dots, m_{|\mathbf{A}|^2}^{(\ell)}$, respectively, and assign the result to $m_{\mathcal{R}}^{(\ell)}$.

Algorithm 1 is derived in Appendix III.

Interestingly, the recursive equations in Corollary 1 do not depend on the time delay τ_k of the signal of user k , i.e. the performance of a CDMA system with multistage detection is independent of the sampling instants and time delays if the assumptions of Corollary 1 on the chip waveforms and on the time delays are satisfied.

Additionally, the dependence of $R^\ell(\lambda)$ on the chip pulse waveforms becomes clear from Algorithm 1: $R^\ell(\lambda)$ depends on $\Phi(\omega)$ through the quantities $\mathcal{E}_s, s = 1, 2, \dots$, defined in (27).

By applying Algorithm 1 we compute the first five asymptotic eigenvalue moments

$$\begin{aligned}
m_{\mathcal{R}}^{(1)} &= \frac{r}{T_c} m_{|\mathbf{A}|^2}^{(1)} \mathcal{E}_1 \\
m_{\mathcal{R}}^{(2)} &= \left(\frac{r}{T_c} \right)^2 [\beta (m_{|\mathbf{A}|^2}^{(1)})^2 \mathcal{E}_2 + m_{|\mathbf{A}|^2}^{(2)} \mathcal{E}_1^2] \\
m_{\mathcal{R}}^{(3)} &= \left(\frac{r}{T_c} \right)^3 [\beta^2 \mathcal{E}_3 (m_{|\mathbf{A}|^2}^{(1)})^3 + 3 m_{|\mathbf{A}|^2}^{(2)} \mathcal{E}_2 \beta m_{|\mathbf{A}|^2}^{(1)} \mathcal{E}_1 + m_{|\mathbf{A}|^2}^{(3)} \mathcal{E}_1^3] \\
m_{\mathcal{R}}^{(4)} &= \left(\frac{r}{T_c} \right)^4 [2\beta^2 \mathcal{E}_2^2 m_{|\mathbf{A}|^2}^{(2)} (m_{|\mathbf{A}|^2}^{(1)})^2 + 4\beta \mathcal{E}_1^2 \mathcal{E}_2 m_{|\mathbf{A}|^2}^{(3)} m_{|\mathbf{A}|^2}^{(1)} + 4\beta^2 \mathcal{E}_1 \mathcal{E}_3 m_{|\mathbf{A}|^2}^{(2)} (m_{|\mathbf{A}|^2}^{(2)})^2 + \beta^3 \mathcal{E}_4 (m_{|\mathbf{A}|^2}^{(1)})^4 \\
&\quad + 2\beta \mathcal{E}_1^2 \mathcal{E}_2 (m_{|\mathbf{A}|^2}^{(2)})^2 + \mathcal{E}_1^4 m_{|\mathbf{A}|^2}^{(4)}] \\
m_{\mathcal{R}}^{(5)} &= \left(\frac{r}{T_c} \right)^5 [m_{|\mathbf{A}|^2}^{(5)} \mathcal{E}_5 \beta^4 + \mathcal{E}_1^5 (m_{|\mathbf{A}|^2}^{(1)})^5 + 5\beta^3 \mathcal{E}_1 \mathcal{E}_4 m_{|\mathbf{A}|^2}^{(2)} (m_{|\mathbf{A}|^2}^{(1)})^3 + 5\beta^3 \mathcal{E}_3 \mathcal{E}_2 m_{|\mathbf{A}|^2}^{(2)} (m_{|\mathbf{A}|^2}^{(1)})^3 \\
&\quad + 5\beta^2 \mathcal{E}_3 \mathcal{E}_1^2 m_{|\mathbf{A}|^2}^{(3)} (m_{|\mathbf{A}|^2}^{(1)})^2 + 5\beta^2 \mathcal{E}_1^2 \mathcal{E}_3 (m_{|\mathbf{A}|^2}^{(2)})^2 m_{|\mathbf{A}|^2}^{(1)} + 5\beta^2 \mathcal{E}_1 \mathcal{E}_2^2 (m_{|\mathbf{A}|^2}^{(2)})^2 m_{|\mathbf{A}|^2}^{(1)} \\
&\quad + 5\beta^2 \mathcal{E}_2^2 \mathcal{E}_1 m_{|\mathbf{A}|^2}^{(3)} (m_{|\mathbf{A}|^2}^{(1)})^2 + 5\beta \mathcal{E}_2 \mathcal{E}_1^3 m_{|\mathbf{A}|^2}^{(4)} m_{|\mathbf{A}|^2}^{(1)} + 5\mathcal{E}_2 \mathcal{E}_1^3 m_{|\mathbf{A}|^2}^{(3)} m_{|\mathbf{A}|^2}^{(2)}].
\end{aligned}$$

In general, the eigenvalue moments of \mathcal{R} depend only on the system load β , the sampling rate $\frac{r}{T_c}$, the eigenvalue distribution of the matrix $\mathbf{A}^H \mathbf{A}$, and \mathcal{E}_s , $s \in \mathbb{Z}^+$. The latter coefficients take into account the effects of the shape of the chip pulse or, equivalently, of the frequency spectrum of the function $\phi(t)$. The asymptotic limits of the diagonal elements of the matrix \mathcal{R}^ℓ corresponding to user k depends also on $|a_k|^2$ but not on the time delay τ_k .

In the special case of chip pulse waveforms $\psi(t)$ having bandwidth not greater than the half of the chip rate, i.e. $B \leq \frac{1}{2T_c}$ the result of Corollary 1 holds for any sets of time delays included synchronous systems.

In Theorem 2, chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$ are considered and the diagonal elements of \mathcal{R}^s are shown to be independent of the time delays of the active users.

Theorem 2 *Let the definitions of Theorem 1 hold.*

We assume that the function $\Phi(\omega)$ is bounded in absolute value and has support $\mathcal{S} \subseteq \left[-\frac{\pi}{T_c}, \frac{\pi}{T_c}\right]$. The vectors \mathbf{s}_k are independent with i.i.d. Gaussian elements $s_{nk} \in \mathbb{C}$ such that $\mathbb{E}\{s_{nk}\} = 0$ and $\mathbb{E}\{|s_{nk}|^2\} = \frac{1}{N}$. Furthermore, the elements a_k of the matrix \mathbf{A} are uniformly bounded for any K . The sequence of the empirical distributions $F_{|\mathbf{A}|^2}^{(K)}(\lambda) = \frac{1}{K} \sum_{k=1}^K 1(\lambda - |a_k|^2)$ converges in law almost surely, as $K \rightarrow \infty$, to a non-random distribution function $F_{|\mathbf{A}|^2}(\lambda)$.

Then, given $|a_k|^2$, the n -th diagonal element of the matrix \mathcal{R}^ℓ , with $n \bmod K = k$, converges almost

surely to a deterministic value, conditionally on $|a_k|^2$,

$$\lim_{K=\beta N \rightarrow \infty} (\mathcal{R}^\ell)_{k,m} = \lim_{K=\beta N \rightarrow \infty} \mathbf{u}_k^{(m)H} \mathbf{T}^{\ell-1} \mathbf{u}_k^{(m)} \stackrel{a.s.}{=} R_\ell(|a_k|^2)$$

with $R_\ell(|a_k|^2)$ determined by the following recursion

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} \lambda R_s(\lambda) \nu_{\ell-s-1} \quad (28)$$

and

$$T_\ell(\omega) = \frac{r}{T_c} \sum_{s=0}^{\ell-1} \beta f(R_{\ell-s-1}) \frac{1}{T_c} |\Phi(\omega)|^2 T_s(\omega) \quad \omega \in \mathcal{S} \quad (29)$$

$$f(R_\ell) = \int \lambda R_\ell(\lambda) dF_{|\mathbf{A}|^2}(\lambda) \quad (30)$$

$$\nu_\ell = \frac{r^2}{2\pi T_c} \int_{\mathcal{S}} |\Phi(\omega)|^2 T_\ell(\omega) d\omega. \quad (31)$$

The recursion is initialized by setting $T_0(\omega) = \frac{T_c}{r}$ and $R_0(\lambda) = 1$.

Theorem 2 is shown in Appendix IV. It applies to Front-end A but, in general, not to Front-end B since Front-end B implies the use of root Nyquist pulses. It is straightforward to verify that Algorithm 1 can be applied to determine $R_\ell(\lambda)$, the asymptotic limit of the diagonal elements and the eigenvalue moments of matrices \mathcal{R} satisfying the conditions of Theorem 2.

The mathematical results presented in this section have important implications on the design and analysis of asynchronous CDMA systems and linear detectors for asynchronous CDMA systems. We elaborate on them in the following section.

VI. EFFECTS OF ASYNCHRONISM, CHIP PULSE WAVEFORMS, AND SETS OF OBSERVABLES

The theoretical framework developed in Section V enables the analysis and design of linear multistage detectors for CDMA systems using optimum and suboptimum statistics and possibly non ideal chip pulse waveforms. In this section we focus on the following aspects:

- 1) Analysis of the effects of chip pulse waveforms and time delay distributions when the multistage detectors are fed by sufficient statistics.
- 2) Impact of the use of sufficient and suboptimum statistics on the complexity and the performance of multistage detectors.

A. Sufficient Statistics

Sufficient statistics impaired by discrete additive Gaussian noise are obtained as output of detector Type A. For chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$ and any set of time delays, Theorem 2 applies. For $B > \frac{1}{2T_c}$ and uniform time delay distribution, Corollary 1 holds. In both cases, as $K, N \rightarrow \infty$ with constant ratio the diagonal elements of the matrix \mathcal{R}^ℓ and the eigenvalue moments $m_{\mathcal{R}}^{(\ell)}$ can be obtained from Algorithm 1. As a consequence of (18), the performance of the large class of multiuser detectors that admit a representation as multistage detectors depends only on the diagonal elements \mathcal{R}^ℓ and the variance of the noise. In large CDMA systems, the SINR depends on the system load β , the sampling rate $\frac{r}{T_c}$, the limit distribution of the received powers $F_{|A|^2}(\lambda)$, the variance of the noise σ^2 , the coefficients \mathcal{E}_ℓ , $\ell \in \mathbb{Z}^+$ and the received powers $|a_k|^2$, but it is independent of the time delay τ_k , in general. For $B \leq \frac{1}{2T_c}$, the SINR is also independent of the time delay distribution. Therefore we can state the following corollary.

Corollary 2 *If the bandwidth of the chip pulse waveform satisfies the constraint $B \leq \frac{1}{2T_c}$, large synchronous and asynchronous CDMA systems have the same performance in terms of SINR when a linear detector that admits a representation as multistage detector is used at the receiver.*

If the time delays and the received amplitudes of the signals are known at the receiver and the sampling rate satisfies the conditions of the sampling theorem, synchronous and asynchronous CDMA systems have the same performance. In [24] is established the equivalence between synchronous and asynchronous CDMA systems using an ideal Nyquist sinc waveform ($B = \frac{1}{2T_c}$) and linear MMSE detector. Corollary 2 generalizes that equivalence to any kind of chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$ and any linear multiuser detector with a multistage representation.

By inspection of Algorithm 1 we can verify that the dependence of $R_\ell(|a_k|^2)$ and $m_{\mathcal{R}}^{(\ell)}$ on the sampling rate $\frac{r}{T_c}$ can be expressed by the following relations

$$R_\ell(|a_k|^2) = \left(\frac{r}{T_c}\right)^\ell R_\ell^*(|a_k|^2) \quad (32)$$

and

$$m_{\mathcal{R}}^{(\ell)} = \left(\frac{r}{T_c}\right)^\ell m_{\mathcal{R}}^{*(\ell)} \quad (33)$$

where $R_\ell^*(|a_k|^2)$ and $m_{\mathcal{R}}^{*(\ell)}$ are independent of the sampling rate $\frac{r}{T_c}$. Thanks to this particular dependence and the fact that $\sigma^2 = \frac{r}{T_c} N_0$, the quadratic forms appearing in (18) when specialized to multistage Wiener filters and in to polynomial expansion detectors, $\xi_{k,m}^H \Xi_{k,m}^{-1} \xi_{k,m}$, $\xi_{k,m}^H \Xi^{-1} \xi$, and $\xi^H \Xi^{-1} \Xi_{k,m} \Xi^{-1} \xi$, are independent of the sampling rate for large systems. Thus, the large system performance of (1) linear multistage detectors optimum in a mean square sense (see (19)), (2) of the polynomial expansion detectors and (3) the matched filters is independent of the sampling rate. This property is not general. Detectors that are not designed to benefit at the best from the available sufficient statistics may improve their performance using different sets of sufficient statistics. Therefore, the large system performance of other multistage detectors like PIC detectors depends on the sampling rate and can eventually improve by increasing the oversampling factor r .

Given a positive real γ , let us consider the chip pulse

$$\Phi(\omega) = \begin{cases} \sqrt{\frac{T_c}{\gamma}} & \text{for } |\omega| \leq \frac{\pi\gamma}{T_c}, \\ 0 & \text{otherwise.} \end{cases} \quad (34)$$

corresponding to a sinc waveform with bandwidth $B = \frac{\gamma}{2T_c}$ and unit energy. For waveform (34) with $\gamma = 1$, $T_c = 1$, and $r = 1$ Algorithm 1 reduces to Algorithm 1 in [18] for synchronous systems. Let us denote by $R_\ell^{(\text{syn})}(|a_k|^2, \beta)$ and $m_{\mathcal{R}^{(\text{syn})}}^{(\ell)}(\beta)$ the values of $R_\ell(|a_k|^2)$ and $m_{\mathcal{R}}^{(\ell)}$ for such a synchronous case and system load β . Then, in general, for chip pulse waveform (34) Algorithm 1 yields

$$R_\ell^{(\text{sinc})}(|a_k|^2) = \left(\frac{r}{T_c}\right)^\ell \mathcal{R}_\ell^{(\text{syn})} \left(|a_k|^2, \frac{\beta}{\gamma}\right) \quad (35)$$

and

$$m_{\mathcal{R}^{(\text{sinc})}}^{(\ell)} = \left(\frac{r}{T_c}\right)^\ell m_{\mathcal{R}^{(\text{syn})}}^{(\ell)} \left(\frac{\beta}{\gamma}\right). \quad (36)$$

Therefore, the same property pointed out in part I of this paper [1] for linear MMSE detectors holds for several multistage detectors (namely, multistage Wiener filters, polynomial expansion detectors, matched filters): In a large asynchronous CDMA system using a sinc function with bandwidth $\frac{\gamma}{2T_c}$ as chip pulse waveform and system load β any multistage detector whose performance is independent of the sampling rate performs as well as in a large synchronous CDMA system with modulation based on root Nyquist chip pulses and system load $\beta' = \frac{\beta}{\gamma}$.

The comparison of synchronous and asynchronous systems with equal chip pulse waveforms enables us to analyze the effects on the system performance of the chip pulse waveforms jointly with the effects of the distribution of time delays. We elaborate on these aspects focusing on root raised cosine chip-pulse

waveforms with roll-off $\vartheta \in [0, 1]$ and on chip pulse waveforms (34) with $\gamma \in [1, 2]$. To simplify the notation, we assume $T_c = 1$. Let

$$S(\omega) = \begin{cases} 1 & 0 \leq |\omega| \leq \pi(1 - \vartheta) \\ \frac{1}{2} \left(1 - \sin \left(\frac{|\omega| - \pi}{2\vartheta} \right) \right) & \pi(1 - \vartheta) \leq |\omega| \leq \pi(1 + \vartheta) \\ 0 & |\omega| \geq \pi(1 + \vartheta). \end{cases}$$

The energy frequency spectrum of a root raised cosine waveform with unit energy is given by $|\Psi_{\text{sqr}}(\omega)|^2 = S(\omega)$. The large system analysis of an asynchronous CDMA system using root raised cosine chip pulse waveform is obtained applying Algorithm 1. The corresponding coefficients $\mathcal{E}_{\text{sqr},s}$, $s = \mathbb{Z}^+$, are given by

$$\mathcal{E}_{\text{sqr},s} = 2^s(1 - \gamma) + \frac{1}{\pi} \int_{\pi(1-\gamma)}^{\pi(1+\gamma)} \sin^s \left(\frac{1}{2\gamma} (\pi - \omega) \right) d\omega.$$

It is well known that in a synchronous CDMA system the performance is maximized using root Nyquist waveforms. In this case the performance is independent of the specific waveform and the bandwidth. It equals the performance of a large synchronous system using the sinc function with bandwidth $\frac{1}{2T_c}$ as chip pulse. Since the root raised cosine pulses are root Nyquist waveforms, they attain the maximum SINR in synchronous systems. The large system performance of multistage Wiener filters for synchronous CDMA systems with a root raised cosine waveform is obtained making use of (19) and Algorithm 1 with $r = 1$ and $\mathcal{E}_s = 1$, $s \in \mathbb{Z}^+$.

In general, chip pulse waveform (34) is not a root Nyquist waveform. For this reason the performance analysis of linear multistage Wiener filters for synchronous CDMA systems [14], [18] is not applicable. In this case characterized by interchip interference we can still apply Theorem 1, sampling at rate $\frac{2}{T_c}$ and assuming a Dirac function $f_T(\tau) = \delta(\tau)$ as probability density function of the time delays. For the chip pulse waveform (34), the matrix $\mathbf{Q}(\Omega) = \mathbf{\Delta}_{\Phi,2}(\Omega, 0) \mathbf{\Delta}_{\Phi,2}^H(\Omega, 0)$ used in the recursion of Theorem 1 is given by

$$\mathbf{Q}(\Omega) = \begin{cases} \frac{1}{\gamma} \begin{pmatrix} 1 & e^{-j\frac{\Omega}{2}} \\ e^{j\frac{\Omega}{2}} & 1 \end{pmatrix} & |\Omega| \leq 2\pi \left(1 - \frac{\gamma}{2} \right) \\ \frac{1}{\gamma} \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} & 2\pi \left(1 - \frac{\gamma}{2} \right) \leq |\Omega| \leq \pi. \end{cases}$$

The large system analysis in the asynchronous case with chip pulse (34) can be readily performed making use of (19) and (35).

In Figure 2 the large system SINR at the output of a multistage Wiener filter with $L = 4$ is plotted as a function of the bandwidth for synchronous and asynchronous CDMA systems based on modulation by root raised cosine or by pulse (34). We assume perfect power control, i.e. $\mathbf{A} = \mathbf{I}$, system load $\beta = 0.5$, and input SNR = 10 dB.

It is well known from theory of synchronous CDMA that interchip interference colors the discrete-time spectrum of the signal and degrades performance. Consistently with that Figure 2 shows that for synchronous CDMA root raised cosine pulses, since they avoid interchip interference, outperform sinc pulses with non-integer ratios of bandwidth to chip rate. Asynchronous CDMA systems with both chip pulse waveforms widely outperform the corresponding synchronous systems. In contrast to the synchronous case, sinc pulses exploit the additional degrees of freedom introduced by increasing the bandwidth better than root raised cosine pulses, since they do not color the spectrum in continuous time domain. Thus, an asynchronous CDMA system with sinc pulses considerably outperforms a system using root raised cosine pulses. Note that for asynchronous systems, the spectral shape in continuous time is relevant, while for synchronous systems the spectral shape in discrete time matters. In both cases the spectrum should be as white as possible to achieve high performance. For asynchronous systems, the spectrum is the less colored, the closer the delay distribution resembles an (eventually discrete) uniform distribution.

In Figure 3 the SINR at the output of a multistage Wiener filter with $L = 8$ is plotted as a function of the system load, parametric in the bandwidth, for SNR = 10 dB. The improvement achievable by asynchronous systems over synchronous systems increases as the the system load increases.

B. Chip Rate Sampling

Chip rate sampling is a widely used approach to generate statistics for asynchronous CDMA systems. It implies the use of root Nyquist chip pulses and makes use of front end Type B. Hereafter, we refer to these CDMA systems as systems B, while we refer to the systems that use sufficient statistics from a front end Type A as systems A.

A bound on the performance of systems B with linear MMSE detectors is in [25]. The performance analysis of linear multistage detectors as $K, N \rightarrow \infty$ with $\frac{K}{N} \rightarrow \beta$ can be performed applying Theorem 1 to the chip pulse waveform at the output of the chip matched filter $\Phi(\omega) = \frac{1}{\sqrt{E_\psi}} |\Psi(\omega)|^2$ and assuming $r = 1$. In order to elaborate further on systems B we focus on the square root raised cosine chip pulse with roll-off

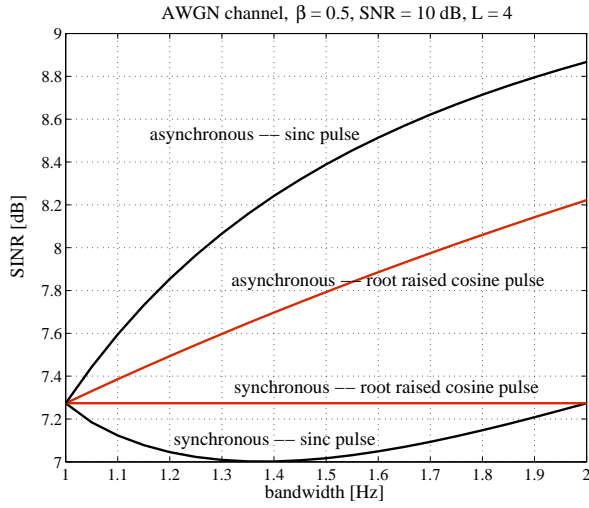


Fig. 2. Output SINR of a multistage Wiener filter with $L = 4$ versus bandwidth. CDMA systems with equal received powers, root raised cosine chip waveforms or sinc pulses, system load $\beta = \frac{1}{2}$ and input SNR = 10 dB are considered.

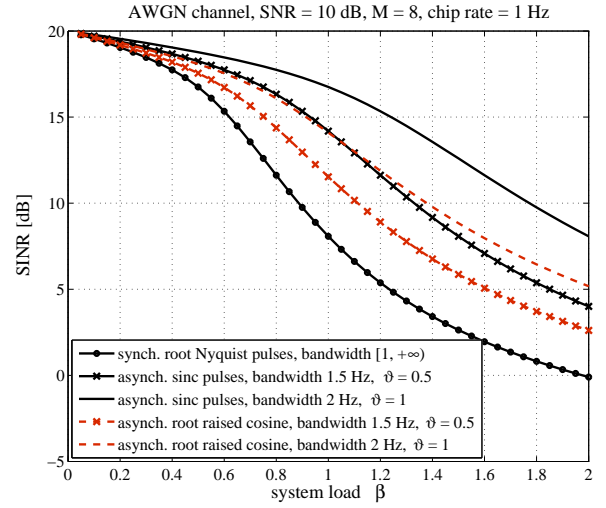


Fig. 3. Output SINR of a multistage Wiener filter with $L = 8$ versus the system load. Asynchronous CDMA systems with equal received powers, root raised cosine chip waveforms or sinc pulses with bandwidth $B = 1.5, 2$ Hz, input SNR = 10 dB are compared to synchronous CDMA systems with root Nyquist chip pulses.

θ [26]

$$\psi(t) = \frac{4\theta(\frac{t}{T_c}) \cos(\pi(1+\theta)\frac{t}{T_c}) + \sin(\pi(1-\theta)\frac{t}{T_c})}{\pi t(1 - (4\theta\frac{t}{T_c})^2)} \quad \theta \in [0, 1]. \quad (37)$$

In this case, the matrix function $\mathbf{Q}(\Omega, \tau) = \mathbf{\Delta}_{\phi,1}(\Omega, \tau) \mathbf{\Delta}_{\phi,1}^H(\Omega, \tau)$ occurring in Theorem 1 reduces to the scalar function

$$\mathbf{Q}(\Omega, \tau) = \begin{cases} \frac{1}{2} + \frac{1}{2} \sin^2\left(\frac{1}{2\theta}(\Omega + \pi)\right) + \frac{\cos 2\pi\tau}{2} \left(1 - \sin^2\left(\frac{1}{2\theta}(\Omega + \pi)\right)\right) & -\pi \leq \Omega \leq -\pi(1-\theta) \\ 1 & -\pi(1-\theta) \leq \Omega \leq \pi(1-\theta) \\ \frac{1}{2} + \frac{1}{2} \sin^2\left(\frac{1}{2\theta}(\Omega - \pi)\right) + \frac{\cos 2\pi\tau}{2} \left(1 - \sin^2\left(\frac{1}{2\theta}(\Omega - \pi)\right)\right) & \pi(1-\theta) \leq \Omega \leq \pi. \end{cases}$$

due to the fact that $r = 1$. Equal received powers, system load $\beta = \frac{1}{2}$, multistage Wiener filters with $L = 3$ define the scenario we consider for the asymptotic analysis.

The analysis shows a strong dependence of the performance on the time delays. As expected, it is possible to verify that the best SINR is obtained when the sampling instants coincide with the time delays of the user of interest.

In Figure 4 we compare the performance of system B with root raised cosine chip pulse to the SINR of a system A with the same modulating pulse. In the comparison we consider the best SINR for system B obtained when the sampling times coincide with the time delays of the user of interest. The curves represent the

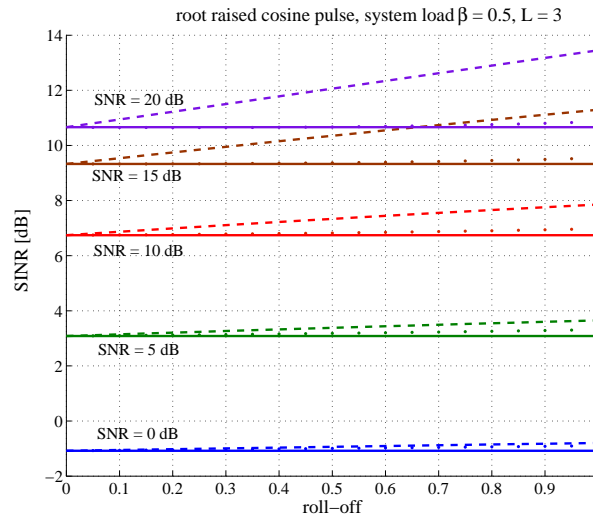


Fig. 4. Asymptotic output SINR of a multistage Wiener filter with $L = 3$ versus the roll-off θ as front-end A (dashed lines) and front-end B (dots) are in use in an asynchronous CDMA system. The solid lines show the reference performance in synchronous CDMA systems. The curves are parametric in the input SNR with SNR varying between 0 dB and 20 dB in steps of 5 dB.

output SINR as a function of the roll-off θ parameterized with respect to SNR. The parameter (SNR) varies from 0 dB to 20 dB in steps of 5 dB. As reference we also plot the performance of synchronous CDMA systems. As expected, multistage detectors with front-end A outperform the corresponding multistage detectors with front-end B.

Interestingly, while linear multistage detectors and asynchronism in system A can compensate to some extent for the loss in spectral efficiency caused by the increasing roll-off and typical of synchronous CDMA systems such a compensation is not possible in systems B. Systems B behave similarly to synchronous CDMA systems. In fact, the SINR for system B is very close to the performance of synchronous systems for any SNR level.

A thorough explanation of these properties based on general analytical results is in Part I Section V [1]. We recapitulate the main idea briefly here. The performance of a large asynchronous CDMA system is governed by an $r \times r$ matrix function in the frequency domain (eq. (24) in [1])⁴. To give an intuition, the system is then equivalent to a MIMO system with r transmit and r receive antennas. The structure of this matrix is such that the matrix is necessarily rank one for synchronous CDMA systems. Thus, only one dimension of the signal space is spanned. On the contrary, for arbitrary delay distributions, i.e. in general for asynchronous systems, the rank of the MIMO system can be higher, eventually, up to r . This

⁴Note that the matrices $T_\ell(\Omega)$ in Theorem 1 can be interpreted as expansion coefficients of this matrix.

implies that asynchronous systems span more of the available dimensions of the signal space resulting in better exploitation of it. When the received signal is sampled at the chip rate, as in the case of Front-end B, and $r = 1$ the processed signal for an asynchronous system only spans a single dimension, just like in synchronous systems, and the performances of synchronous and asynchronous systems are very similar.

Since the SINR in system B heavily depends on the sampling instants with respect to τ_k , different statistics are needed for the detection of different users in order to obtain good performance. As consequence, joint detection is not feasible and each user has to be detected independently. This is a significant drawback when several or all users have to be detected (e.g. uplink) and has a relevant impact on the complexity of the system. For example, the complexity order per bit of a multistage Wiener filter or polynomial expansion detector is linear in rK in system A while the complexity order per bit of the same detectors is quadratic in K in system B. A similar increase in complexity can be noticed also for other detectors (e.g. linear MMSE detectors, or any multistage detector).

VII. CONCLUSIONS

In Part II of this work we provided guidelines for the design of asynchronous CDMA systems via the analysis of the effects of chip pulse waveforms, time delay distributions, sufficient and suboptimum observables on the complexity and performance of the broad class of multiuser detectors with multistage representation.

Similarly to the results obtained in part I of this article [1], i.e. the chip-pulse constrained spectral efficiency and the performance of linear MMSE detectors, multistage detectors show performance independent of the time delays of the active users if the bandwidth of the chip pulse waveform is not greater than half of the chip rate, i.e. $B \leq \frac{1}{2T_c}$. Above that threshold the performances of linear multistage detectors depend on the time delay distributions and asynchronous CDMA systems outperform synchronous CDMA systems.

The framework presented here enabled the analysis of optimum and suboptimum multistage detectors based on front ends whose sampled noise outputs are white. We focused on multistage detectors using statistics (A), which are sufficient, or observables (B), which are suboptimum. In the two cases of (i) chip pulses with bandwidth $B \leq \frac{1}{2T_c}$ and (ii) chip pulses with bandwidth $B > \frac{1}{2T_c}$, sufficient statistics, and uniform distribution, the effects of the chip pulse waveforms on the detector performance are described by the coefficients $\mathcal{E}_s = \frac{1}{2\pi T_c^{s-1}} \int_{-2\pi B}^{2\pi B} |\Psi(\omega)|^{2s} d\omega$. The output SINR of linear MMSE detectors, multistage Wiener filters, polynomial expansion detectors, and matched filters is independent of the sampling rate. In contrast, the output SINR of other multistage detectors like PIC detectors depends on the sampling rate and

increases with it.

Comparing the performance of synchronous and asynchronous CDMA systems with modulation based on root Nyquist pulses, namely root raised cosine waveforms, and modulation based on sinc functions with increasing bandwidth, it becomes apparent that the chip pulse design for synchronous CDMA systems follows the same guidelines as the chip pulse design for single user systems. In contrast, chip pulse design for asynchronous CDMA systems is governed by entirely different rules. In fact, for example, we found that CDMA systems with uniform delay distributions perform well if the spectrum of the received signal is as white as possible.

The asymptotic analysis of asynchronous CDMA systems using statistics (B) shows that the performance of multistage Wiener filters is close to the SINR of the corresponding synchronous CDMA systems for any bandwidth and level of SNR. Therefore, this kind of front-end is not capable of exploiting the benefits of asynchronous CDMA.

The universal weights proposed for the design of low complexity detectors account for the effects of asynchronism, sub-optimality of the statistics, and non-ideality of pulse-shapers. They depend on the sampling rate although the large system performance do not.

From the asymptotic analysis and design performed in this work we can draw the following conclusions:

- Multistage detectors with front end Type B and universal weights are asymptotically suboptimal and have the same complexity order per bit $\mathcal{O}(K^2)$ in uplink as the linear MMSE detector.
- Multistage Wiener filters and polynomial expansion detectors with statistics A and universal weights are asymptotically optimum and have the same complexity order per bit as the matched filter, i.e. $\mathcal{O}(rK)$ with $r \ll K$.
- If only a user has to be detected, multistage detectors using statistics (B) have slightly lower complexity than multistage detectors with statistics (A), namely they have a complexity per bit $\mathcal{O}(K^2)$ while in the later case the complexity per bit is $\mathcal{O}(rK^2)$. However, they perform almost as the multistage detectors for synchronous systems at any SNR and do not provide the gain in performance due to asynchronism in contrast to statistics (A).

ACKNOWLEDGMENT

The authors thank Alex Grant and Dirk Slock for useful discussions.

APPENDIX I

PROOF OF THEOREM 1

Before going into the details of the proof we introduce some properties of the convergence in probability and in probability one.

Property A: Let us consider a finite number q of random sequences $\{a_n^{(1)}\}, \dots, \{a_n^{(q)}\}$ that converge in probability to deterministic limits a_1, \dots, a_q , respectively. Then, any linear combination of such sequences converges in probability to the linear combination of the limits. Furthermore, if $|a_n^{(s)} - a_s| \xrightarrow{\mathcal{P}} o(N^{-i_s})$, with $i_s \in \mathbb{R}^+$, and $s = 1, \dots, q$, then any linear combination of the random sequences converges as $o(N^{-\min_{s=1, \dots, q}(i_s)})$, at worst.

Property B: Let $\{a_n\}$ and $\{b_n\}$ be two random sequences that converge in probability to a and b , respectively. Then, the sequence $\{a_n b_n\}$ converges in probability to ab .

Property C: If for large n , $\Pr\{|a_n - a| > \varepsilon\} \leq o(n^{-s})$ and $\Pr\{|b_n - b| > \varepsilon\} \leq o(n^{-t})$, with $s, t \in \mathbb{R}^+$, then also $\Pr\{|(a_n - a)(b_n - b)| > \varepsilon\} \leq o(n^{-\min(s, t)})$, at worst.

The convergence in probability one or almost sure convergence implies the convergence in probability. In general, the converse is not true. However, if a random sequence a_k converge in probability to a constant a with a convergence rate $o(n^{-s})$ and $s > 1$, i.e. $\Pr\{|a_n - a| > \varepsilon\} \leq o(n^{-s})$, then, also the convergence in probability one holds. This is a straightforward consequence of the Borel Cantelli lemma (see e.g. [27]).

In part I Theorem 3 of this work [1] we have shown that, when $K, N \rightarrow +\infty$ with constant ratio β , the eigenvalue distribution of the infinite matrix \mathcal{R} is the same as the eigenvalue distribution of the matrix $\tilde{\mathbf{R}} = \mathbf{A}^H \tilde{\mathbf{S}}^H \tilde{\mathbf{S}} \mathbf{A} = \tilde{\mathbf{H}}^H \tilde{\mathbf{H}}$ where $\tilde{\mathbf{S}} = (\tilde{\Phi}_1 \mathbf{s}_1, \tilde{\Phi}_2 \mathbf{s}_2, \dots, \tilde{\Phi}_K \mathbf{s}_K)$ and $\tilde{\Phi}_k$ is the r -block-wise circulant matrix of order N defined in (10) with $\tilde{\tau}_k = \tau_k \bmod T_c$.

Let us consider the block diagonal matrix $\Delta_{\phi, r}(\tilde{\tau}_k)$ with $r \times 1$ blocks

$$(\Delta_{\phi, r}(\tilde{\tau}_k))_{\ell, \ell} = \begin{pmatrix} \phi\left(2\pi \frac{\ell-1}{N}, \tilde{\tau}_k\right) \\ \phi\left(2\pi \frac{\ell-1}{N}, \tilde{\tau}_k - \frac{T_c}{r}\right) \\ \vdots \\ \phi\left(2\pi \frac{\ell-1}{N}, \tilde{\tau}_k - \frac{r-1}{r}T_c\right) \end{pmatrix}. \quad (38)$$

and introduce the matrices

$$\hat{\mathbf{S}} = (\Delta_{\phi, r}(\tilde{\tau}_1) \mathbf{s}_1, \Delta_{\phi, r}(\tilde{\tau}_2) \mathbf{s}_2, \dots, \Delta_{\phi, r}(\tilde{\tau}_K) \mathbf{s}_K) \quad (39)$$

and $\hat{\mathbf{R}} = \mathbf{A}^H \hat{\mathbf{S}}^H \hat{\mathbf{S}} \mathbf{A}$.

By applying the same approach as in part I Theorem 1 of this work [1] it can be shown that the eigenvalue distribution of the matrices $\tilde{\mathbf{R}}$ and $\hat{\mathbf{R}}$ coincide. Then, also the eigenvalue moments of the two matrices coincide. The same property holds for the diagonal elements of the matrices $\tilde{\mathbf{R}}^\ell$ and $\hat{\mathbf{R}}^\ell$ with $\ell \in \mathbb{Z}^+$.

In the following we focus on the asymptotic analysis of the diagonal elements of the matrices $\hat{\mathbf{R}}^\ell$.

Throughout this proof we adopt the following notation. For $k = 1, \dots, K$ and $n = 1, \dots, N$

- $\hat{\mathbf{h}}_k$ is the k^{th} column of the matrix $\widehat{\mathbf{H}}$;
- $\hat{\mathbf{h}}_{nk}$ is the n^{th} $r \times 1$ block of the vector $\hat{\mathbf{h}}_k$ and $\hat{\mathbf{h}}_{nk} = a_k(\Delta_{\phi,r}(\tilde{\tau}_k))_{nn} s_{nk}$;
- $\hat{\delta}_n$ is the n^{th} block row of $\widehat{\mathbf{H}}$ of dimensions $r \times K$;
- $\widehat{\mathbf{H}}_{\models n}$ is the matrix obtained from $\widehat{\mathbf{H}}$ by suppressing $\hat{\delta}_n$;
- $\widehat{\mathbf{H}}_{\sim k}$ is the matrix obtained from $\widehat{\mathbf{H}}$ by suppressing $\hat{\mathbf{h}}_k$;
- $\hat{\mathbf{T}} = \widehat{\mathbf{H}}\widehat{\mathbf{H}}^H$ and $\hat{\mathbf{T}}_{\sim k} = \widehat{\mathbf{H}}_{\sim k}\widehat{\mathbf{H}}_{\sim k}^H$;
- $\hat{\mathbf{R}}_{\models n} = \widehat{\mathbf{H}}_{\models n}^H \widehat{\mathbf{H}}_{\models n}$;
- $\hat{\boldsymbol{\sigma}}_n = (s_{n1}, s_{n2}, \dots, s_{nK})$;
- $\nabla_{n,t}$, for $t = 1, \dots, r$ and $n = 1, \dots, N$, is a $K \times K$ diagonal matrix with the k^{th} element equal to $\phi\left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{(t-1)T_c}{r}\right)$. Note that $\hat{\boldsymbol{\sigma}}_n \nabla_{n,t} \mathbf{A}$ coincides with the $(t + (n-1)r)^{\text{th}}$ row of the matrix $\widehat{\mathbf{H}}$.
- $\hat{\mathbf{T}}_{[nn]}^s$ is the n^{th} diagonal block of $\hat{\mathbf{T}}^s$ of dimensions $r \times r$.

Furthermore, since the channel gains a_k are bounded, we denote by a_{MAX} their upper bound, i.e. $|a_k| < a_{\text{MAX}}, \forall k$. Finally, thanks to the assumption that $\Phi(\omega)$ is bounded in absolute value with finite support also $\phi(\Omega, \tau)$ is upper bounded for any Ω and τ . We denote by Φ_{MAX} its bound.

Let us observe first that the eigenvalue moments of the matrix $\hat{\mathbf{R}}$ (or equivalently of $\hat{\mathbf{T}}$) are almost surely upper bounded by a finite positive values $C^{(s)}$, i.e.

$$\exists C^{(s)} < +\infty : \quad \Pr \left\{ \frac{1}{N} \text{tr} \hat{\mathbf{R}}^s < C^{(s)} \right\} = 1 \quad \text{as } K, N \rightarrow +\infty, \frac{K}{N} \rightarrow \beta. \quad (40)$$

In fact,

$$\begin{aligned} \frac{1}{N} \text{tr} \hat{\mathbf{R}}^s &= \frac{1}{N} \sum_{k_1, \dots, k_s=1}^K \sum_{n_1, \dots, n_s=1}^N \hat{\mathbf{h}}_{n_1, k_1}^H \hat{\mathbf{h}}_{n_1, k_2} \hat{\mathbf{h}}_{n_2, k_2}^H \hat{\mathbf{h}}_{n_2, k_3} \dots \hat{\mathbf{h}}_{n_s, k_s}^H \hat{\mathbf{h}}_{n_s, k_1} \\ &= \frac{1}{N} \sum_{k_1, \dots, k_s=1}^K |a_{k_1}|^2 \dots |a_{k_s}|^2 \sum_{n_1, \dots, n_s=1}^N \Delta_{\phi,r}(\tilde{\tau}_1)_{n_1 n_1}^H \Delta_{\phi,r}(\tilde{\tau}_2)_{n_1 n_1} \dots \Delta_{\phi,r}(\tilde{\tau}_s)_{n_s n_s}^H \Delta_{\phi,r}(\tilde{\tau}_1)_{n_s n_s} \times \\ &\quad \times s_{n_1, k_1}^* s_{n_1, k_2} s_{n_2, k_2}^* s_{n_2, k_3} \dots s_{n_s, k_s}^* s_{n_s, k_1} \end{aligned}$$

Applying the approach of non-crossing partitions [28], [29], it is possible to recognize that the factors $s_{n_1, k_1}^* s_{n_1, k_2} s_{n_2, k_2}^* s_{n_2, k_3} \dots s_{n_s, k_s}^* s_{n_s, k_1}$ which do not vanish asymptotically, correspond to the ones having

nonzero non-crossing partitions. Correspondingly, also the remaining factors

$$\Delta_{\phi,r}(\tilde{\tau}_1)_{n_1 n_1}^H \Delta_{\phi,r}(\tilde{\tau}_2)_{n_1 n_1} \cdots \Delta_{\phi,r}(\tilde{\tau}_s)_{n_s n_s}^H \Delta_{\phi,r}(\tilde{\tau}_1)_{n_s n_s}$$

are positive and bounded by

$$|\Delta_{\phi,r}(\tilde{\tau}_1)_{n_1 n_1}^H \Delta_{\phi,r}(\tilde{\tau}_2)_{n_1 n_1} \cdots \Delta_{\phi,r}(\tilde{\tau}_s)_{n_s n_s}^H \Delta_{\phi,r}(\tilde{\tau}_1)_{n_s n_s}| \leq \frac{r^{2s} \Delta_{\text{MAX}}^{2s}}{T_c^{2s}}.$$

Therefore,

$$\frac{1}{N} \text{Tr} \hat{\mathbf{R}}^s \leq \frac{r^{2s} \Delta_{\text{MAX}} a_{\text{MAX}}^{2s}}{T_c^{2s}} \left(\frac{1}{N} \sum_{k_1, \dots, k_s=1}^K \sum_{n_1, \dots, n_s=1}^N s_{n_1, k_1}^* s_{n_1, k_2} s_{n_2, k_2}^* s_{n_2, k_3} \cdots s_{n_s, k_s}^* s_{n_s, k_1} \right). \quad (41)$$

The last factor in (41) is the s -th eigenvalue moment of a central Wishart matrix with zeromean i.i.d Gaussian entries having variance $\frac{1}{N}$. Well established results of random matrix theory [30], [29], [12] show that the eigenvalue moments of such a matrix converge almost surely to finite values. More specifically,

$$\frac{1}{N} \sum_{n_1, \dots, n_s=1}^N s_{n_1, k_1}^* s_{n_1, k_2} s_{n_2, k_2}^* s_{n_2, k_3} \cdots s_{n_s, k_s}^* s_{n_s, k_1} \xrightarrow{a.s.} \sum_{i=0}^{s-1} \binom{s}{i} \binom{s}{i+1} \frac{\beta^i}{s}. \quad (42)$$

Then, appealing to (41) and (42), the eigenvalue moments of the matrices $\hat{\mathbf{R}}$ and $\hat{\mathbf{T}}$ are upper bounded almost surely by

$$C^{(s)} = \frac{r^{2s} \Delta_{\text{MAX}} a_{\text{MAX}}^{2s}}{T_c^{2s}} \sum_{i=0}^{s-1} \binom{s}{i} \binom{s}{i+1} \frac{\beta^i}{s}. \quad (43)$$

The proof of Theorem 1 is based on strong induction. In the first step we prove the following facts:

- 1) The diagonal elements of the matrix $\hat{\mathbf{R}}$ converge almost surely, as $N \rightarrow \infty$, to deterministic values $R_1(|a_k|^2, \tilde{\tau}_k)$, conditionally on $(|a_k|^2, \tilde{\tau}_k)$. Furthermore, $\forall \varepsilon > 0$ and large $K = \beta N$

$$\Pr\{|\hat{\mathbf{R}}_{kk} - R_1(|a_k|^2, \tilde{\tau}_k)| > \varepsilon\} \leq o(N^{-2}).$$

- 2) $\hat{\mathbf{T}}_{[nn]}$, the $r \times r$ block diagonal elements of the matrix $\hat{\mathbf{T}} = \hat{\mathbf{H}} \hat{\mathbf{H}}^H$, converge almost surely to deterministic blocks $\mathbf{T}_1(\Omega)$, with $\Omega = \lim_{N \rightarrow \infty} 2\pi \frac{n}{N}$. Additionally, $\forall \varepsilon > 0$, large $K = \beta N$ and $u, v = 1, \dots, r$,

$$\Pr\{|\hat{\mathbf{T}}_{[nn]} - \mathbf{T}_1(\Omega)|_{uv} > \varepsilon\} \leq o(N^{-2}).$$

Then, in the recursion step, we use the following induction assumptions:

- 1) For $s = 1, \dots, \ell - 1$, the diagonal elements of the matrix $\hat{\mathbf{R}}^s$, converge almost surely, as $K = \beta N \rightarrow \infty$, to deterministic values $R_s(|a_k|^2, \tilde{\tau}_k)$, conditionally on $(|a_k|^2, \tilde{\tau}_k)$. Additionally, $\forall \varepsilon > 0$ and large $K = \beta N$, $\Pr\{|\hat{\mathbf{R}}^s_{kk} - R_s(|a_k|^2, \tilde{\tau}_k)| > \varepsilon\} \leq o(N^{-2})$.

- 2) For $s = 1, \dots, \ell - 1$, $\widehat{\mathbf{T}}_{[nn]}^s$, the $r \times r$ block diagonal elements of the matrix $\widehat{\mathbf{T}}^s$ converge almost surely to deterministic blocks $\mathbf{T}_s(\Omega)$, with⁵ $\Omega = \lim_{N \rightarrow \infty} 2\pi \frac{n}{N}$. Additionally, $\forall \varepsilon > 0$, large $K = \beta N$, and $u, v = 1, \dots, r$, $\Pr\{ |(\widehat{\mathbf{T}}_{[nn]}^s)_{uv} - (\mathbf{T}_s(\Omega))_{uv}| > \varepsilon \} \leq o(N^{-2})$.

We prove:

- 1) The diagonal elements of the matrix $\widehat{\mathbf{R}}^\ell$, converge almost surely, as $K = \beta N \rightarrow \infty$, to deterministic values $R^\ell(|a_k|^2, \tilde{\tau}_k)$, conditionally on $(|a_k|^2, \tilde{\tau}_k)$. Furthermore, $\forall \varepsilon > 0$ and large $K = \beta N$

$$\Pr\{ |(\widehat{\mathbf{R}}^\ell)_{kk} - R^\ell(|a_k|^2, \tilde{\tau}_k)| > \varepsilon \} \leq o(N^{-2}). \quad (44)$$

- 2) The blocks $\widehat{\mathbf{T}}_{[nn]}^\ell$, converge almost surely to deterministic blocks $\mathbf{T}^\ell(\Omega)$ with $\lim_{N \rightarrow \infty} 2\pi \frac{n}{N}$. Additionally, $\forall \varepsilon > 0$, large N and $u, v = 1, \dots, r$,

$$\Pr\{ |(\widehat{\mathbf{T}}_{[nn]}^\ell)_{uv} - (\mathbf{T}^\ell(\Omega))_{uv}| > \varepsilon \} \leq o(N^{-2}). \quad (45)$$

First step: Consider $\widehat{\mathbf{R}}_{kk} = \widehat{\mathbf{h}}_k^H \widehat{\mathbf{h}}_k = |a_k|^2 \mathbf{s}_k^H \Delta_{\phi,r}^H(\tilde{\tau}_k) \Delta_{\phi,r}(\tilde{\tau}_k) \mathbf{s}_k$. Thanks to the bound $|\phi(\Omega, \tau)| < \Phi_{\text{MAX}}$ which holds for any Ω and τ , also the eigenvalues of the matrix $\Delta_{\phi,r}^H(\tilde{\tau}) \Delta_{\phi,r}(\tilde{\tau})$ are upper bounded. In fact, they are given by $\sum_{t=1}^r \left| \phi \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{(t-1)T_c}{r} \right) \right|^2$ for $n = 1, \dots, N$. Therefore, the limit eigenvalue distribution of the matrix $\Delta_{\phi,r}^H(\tilde{\tau}) \Delta_{\phi,r}(\tilde{\tau})$ has upper bounded support Δ_{MAX} . Then, by appealing to Lemma 9 in part I [1] with $p = 4$ and by making use of the bound for any Hermitian matrix $\mathbf{C} \in \mathbb{C}^{N \times N}$, $(\text{tr} \mathbf{C})^2 \leq N \text{tr}(\mathbf{C}^2)$ we obtain

$$\begin{aligned} \zeta_1 &= \mathbb{E} \left| |a_k|^2 \mathbf{s}_k^H \Delta_{\phi,r}^H(\tilde{\tau}_k) \Delta_{\phi,r}(\tilde{\tau}_k) \mathbf{s}_k - \frac{|a_k|^2}{N} \text{tr}(\Delta_{\phi,r}^H(\tilde{\tau}_k) \Delta_{\phi,r}(\tilde{\tau}_k)) \right|^4 \\ &\leq \frac{K_4 |a_k|^4}{N^3} \text{tr}(\Delta_{\phi,r}^H(\tilde{\tau}_k) \Delta_{\phi,r}(\tilde{\tau}_k))^4 \\ &\leq \frac{K_4 |a_k|^4}{N^2} \Delta_{\text{MAX}}^4. \end{aligned}$$

Since $|a_k| \leq a_{\text{MAX}} < +\infty$, the Bienaymé inequality yields $\forall \varepsilon > 0$

$$\begin{aligned} \Pr \left\{ \left| \widehat{\mathbf{R}}_{kk} - \frac{|a_k|^2}{N} \text{tr}(\Delta_{\phi,r}^H(\tilde{\tau}_k) \Delta_{\phi,r}(\tilde{\tau}_k)) \right| \geq \varepsilon \right\} &\leq \frac{\mathbb{E} \left| \widehat{\mathbf{R}}_{kk} - \frac{|a_k|^2}{N} \text{tr}(\Delta_{\phi,r}^H(\tilde{\tau}_k) \Delta_{\phi,r}(\tilde{\tau}_k)) \right|^4}{\varepsilon^4} \\ &\leq \frac{K_4 |a_k|^4 \Delta_{\text{MAX}}^4}{N^2 \varepsilon^4} \end{aligned} \quad (46)$$

⁵Note that $n = n(N)$ is also a function of the matrix size N .

Thanks to the bound (46) $\forall \varepsilon > 0$

$$\Pr \left\{ \left| \widehat{\mathbf{R}}_{kk} - R_1(|a_k|^2, \tilde{\tau}_k) \right| \geq \varepsilon \right\} \leq o(N^{-2}).$$

Furthermore, appealing to the Borel Cantelli lemma (see e.g. [27]), this bound implies the following almost sure convergence.

$$\begin{aligned} R_1(\lambda, \tau)|_{(\lambda, \tau)=(|a_k|^2, \tau_k)} &= \lim_{K=\beta N \rightarrow \infty} \widehat{\mathbf{R}}_{kk} \\ &= \lim_{K=\beta N \rightarrow \infty} \frac{|a_k|^2}{N} \text{tr}(\Delta_{\phi, r}^H(\tilde{\tau}_k) \Delta_{\phi, r}(\tilde{\tau}_k)) \\ &= \lim_{K=\beta N \rightarrow \infty} \frac{|a_k|^2}{N} \sum_{\ell=1}^N (\Delta_{\phi, r}^H(\tilde{\tau}_k))_{\ell, \ell} (\Delta_{\phi, r}(\tilde{\tau}_k))_{\ell, \ell} \\ &= \frac{\lambda}{2\pi} \int_0^{2\pi} \Delta_{\phi, r}^H(\Omega, \tau) \Delta_{\phi, r}(x, \tau) dx \Big|_{(\lambda, \tau)=(|a_k|^2, \tilde{\tau}_k)}. \end{aligned} \quad (47)$$

Let us now consider the block matrix $\widehat{\mathbf{T}}_{[nn]}$ whose (u, v) element $(\widehat{\mathbf{T}}_{[nn]})_{uv}$ is given by

$$(\widehat{\mathbf{T}}_{[nn]})_{uv} = \widehat{\boldsymbol{\sigma}}_n \mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H \widehat{\boldsymbol{\sigma}}_n^H.$$

Thanks to the assumption of Theorem 1 that the support of $F_{|\mathbf{A}|^2, T}(\lambda, \tau)$ is bounded and $\phi(\Omega, \tau)$ is bounded in absolute value, the diagonal elements of the diagonal matrix $\mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H$ are upper bounded in absolute value by a positive constant T_{MAX} . Then, by appealing to Lemma 9 in part I [1] we obtain

$$\begin{aligned} \mathbb{E} \left(\left| (\widehat{\mathbf{T}}_{[nn]})_{u,v} - \frac{1}{N} \text{tr}(\mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H) \right|^4 \right) &\leq \frac{K_4}{N^3} \text{tr}(\mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H)^4 \\ &\leq \frac{K_4}{N^2} T_{\text{MAX}}^4. \end{aligned} \quad (48)$$

By appealing again to the Bienaymé inequality and by making use of the bound (48) we obtain $\forall \varepsilon > 0$

$$\begin{aligned} \Pr \left\{ \left| (\widehat{\mathbf{T}}_{[nn]})_{u,v} - \frac{1}{N} \text{tr}(\mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H) \right| > \varepsilon \right\} &\leq \frac{1}{\varepsilon^4} \mathbb{E} \left(\left| (\widehat{\mathbf{T}}_{[nn]})_{u,v} - \frac{1}{N} \text{tr}(\mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H) \right|^4 \right) \\ &\leq \frac{K_4 T_{\text{MAX}}^4}{\varepsilon^4 N^2}. \end{aligned} \quad (49)$$

Thus, the following convergence in probability holds

$$\begin{aligned} \lim_{K=\beta N \rightarrow \infty} (\widehat{\mathbf{T}}_{[nn]})_{u,v} &= \lim_{K=\beta N \rightarrow \infty} \frac{1}{N} \text{tr}(\mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H) \\ &= \lim_{K=\beta N \rightarrow \infty} \frac{\beta}{K} \sum_{k=1}^K |a_k|^2 \phi \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{u-1}{r} T_c \right) \phi^* \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{v-1}{r} T_c \right) \\ &= \beta \int \lambda \phi \left(\Omega, \tau - \frac{u-1}{r} T_c \right) \phi \left(\Omega, \tau - \frac{v-1}{r} T_c \right) dF_{|\mathbf{A}|^2, T}(\lambda, \tau), \end{aligned} \quad (50)$$

with $\Omega = \lim_{N \rightarrow \infty} 2\pi \frac{n}{N}$ and $0 \leq \Omega \leq 2\pi$. Therefore, the block matrix $\widehat{\mathbf{T}}_{[nn]}$ converges in probability and in mean square sense to the $r \times r$ matrix

$$\begin{aligned} \mathbf{T}_1(\Omega) &= \lim_{K=\beta N \rightarrow \infty} \widehat{\mathbf{T}}_{[nn]} \\ &= \beta \int \lambda \Delta_{\phi,r}(\Omega, \tau) \Delta_{\phi,r}^H(\Omega, \tau) dF_{|\mathbf{A}|^2, T}(\lambda, \tau) \end{aligned}$$

with $0 \leq \Omega \leq 2\pi$. Thanks to the bound (48) for large $K = \beta N$ and $\forall \varepsilon > 0$ the bound

$$\Pr \left\{ \left| (\widehat{\mathbf{T}}_{[nn]})_{u,v} - (\mathbf{T}(\Omega))_{u,v} \right| < \varepsilon \right\} \leq o(N^{-2})$$

holds. Making use of this bound and applying the Borel Cantelli lemma the almost sure convergence is also proven. This concludes the proof of the first step.

Step ℓ :

By appealing to the induction assumptions, i.e. the almost sure convergence of the diagonal elements of $\widehat{\mathbf{R}}^s$ and of the diagonal $r \times r$ blocks of $\widehat{\mathbf{T}}^s$, for $s = 1, \dots, \ell - 1$, we prove that the following almost sure convergence holds:

$$\begin{aligned} \lim_{K=\beta N \rightarrow \infty} \frac{\text{tr} \mathbf{A} \nabla_{n,u} \widehat{\mathbf{R}}_{\mathbb{T}n}^s \nabla_{n,v}^H \mathbf{A}^H}{N} &= \lim_{K=\beta N \rightarrow \infty} \sum_{k=1}^K \frac{|a_k|^2}{N} \phi \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{u-1}{r} T_c \right) \phi^* \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{v-1}{r} T_c \right) (\widehat{\mathbf{R}}_{\mathbb{T}n}^s)_{kk} \\ &= \beta \int \lambda \phi \left(\Omega, \tau - \frac{u-1}{r} T_c \right) \phi^* \left(\Omega, \tau - \frac{v-1}{r} T_c \right) R_s(\lambda, \tau) dF_{|\mathbf{A}|^2, T}(\lambda, \tau) \end{aligned} \quad (51)$$

with $\Omega = \lim_{N \rightarrow \infty} 2\pi \frac{n-1}{N}$, $s = 1, \dots, \ell - 1$ and

$$R_s(\lambda, \tau) \Big|_{(\lambda, \tau) = (|a_k|^2, \tilde{\tau}_k)} = \lim_{K=\beta N \rightarrow \infty} (\widehat{\mathbf{R}}^s)_{kk} + o(N^{-2}) \quad (52)$$

as from the recursion assumptions. Furthermore, we prove the following almost sure convergence

$$\begin{aligned} \lim_{K=\beta N \rightarrow \infty} \frac{|a_k|^2}{N} \text{tr} \Delta_{\phi,r}^H(\tilde{\tau}_k) \widehat{\mathbf{T}}_{\sim k}^s \Delta_{\phi,r}(\tilde{\tau}_k) &= \lim_{K=\beta N \rightarrow \infty} \frac{|a_k|^2}{N} \sum_{n=1}^N (\Delta_{\phi,r}^H(\tilde{\tau}_k))_{nn} (\widehat{\mathbf{T}}^s)_{nn} (\Delta_{\phi,r}(\tilde{\tau}_k))_{nn} \\ &= \frac{\lambda}{2\pi} \int_0^{2\pi} \Delta_{\phi,r}^H(\Omega, \tau) \mathbf{T}_s(\Omega) \Delta_{\phi,r}(\Omega, \tau) d\Omega \Big|_{(\lambda, \tau) = (|a_k|^2, \tilde{\tau}_k)} \end{aligned} \quad (53)$$

with $s = 1, \dots, \ell - 1$ and

$$\mathbf{T}_s(\Omega) = \lim_{K=\beta N \rightarrow \infty} (\widehat{\mathbf{T}}^s)_{nn}. \quad (54)$$

In fact, for (51) we can write

$$\begin{aligned}\zeta_2 &= \Pr \left\{ \left| \frac{1}{N} \text{tr} \mathbf{A} \nabla_{n,u} \hat{\mathbf{R}}_{\mathbb{F}n}^s \nabla_{n,v}^H \mathbf{A}^H \right. \right. \\ &\quad \left. \left. - \frac{1}{N} \sum_{k=1}^K |a_k|^2 \phi \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{u-1}{r} T_c \right) \phi^* \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{v-1}{r} T_c \right) R_s(|a_k|^2, \tilde{\tau}_k) \right| > \varepsilon \right\} \\ &\leq \zeta_{2a} + \zeta_{2b}\end{aligned}$$

where

$$\zeta_{2a} = \Pr \left\{ \left| \frac{1}{N} \text{tr} \mathbf{A} \nabla_{n,u} (\hat{\mathbf{R}}^s - \hat{\mathbf{R}}_{\mathbb{F}n}^s) \nabla_{n,v}^H \mathbf{A}^H \right| > \frac{\varepsilon}{2} \right\}$$

and

$$\zeta_{2b} = \Pr \left\{ \left| \frac{1}{N} \sum_{k=1}^K |a_k|^2 \phi \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{u-1}{r} T_c \right) \phi^* \left(2\pi \frac{n-1}{N}, \tilde{\tau}_k - \frac{v-1}{r} T_c \right) \left((\hat{\mathbf{R}}^s)_{kk} - R_s(|a_k|^2, \tilde{\tau}_k) \right) \right| > \frac{\varepsilon}{2} \right\}.$$

Note that

$$\zeta_{2a} \leq \Pr \left\{ \left| \frac{1}{K} \text{tr} (\hat{\mathbf{R}}^s - \hat{\mathbf{R}}_{\mathbb{F}n}^s) \right| > \frac{\varepsilon}{2\beta a_{\text{MAX}}^2 \phi_{\text{MAX}}^2} \right\}.$$

The expansion of the matrix $\hat{\mathbf{R}}^s = (\hat{\mathbf{R}}_{\mathbb{F}n} + \hat{\boldsymbol{\delta}}_n^H \hat{\boldsymbol{\delta}}_n)^s$ yields

$$\text{tr} \hat{\mathbf{R}}^s = \text{tr} \hat{\mathbf{R}}_{\mathbb{F}n}^s + \sum_{\substack{(i_0, i_1, \dots, i_{s-1}) \\ i_0 + \sum_{j=1}^{s-1} (j+1) i_j = s_0}} \varphi(i_0, i_1, \dots, i_{s-1}) \prod_{u=0}^{s-1} \left(\hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\mathbb{F}n}^u \hat{\boldsymbol{\delta}}_n \right)^{i_u}$$

where $\varphi(i_0, i_1, \dots, i_{s-1}) \leq 2^s$ is the number of the terms of the expansion of $\hat{\mathbf{R}}^s$ whose trace equals $\prod_{u=0}^{s-1} \left(\hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\mathbb{F}n}^u \hat{\boldsymbol{\delta}}_n \right)^{i_u}$. Then,

$$\zeta_{2a} \leq 2^s \sum_{\substack{(i_0, i_1, \dots, i_{s-1}) \\ i_0 + \sum_{j=1}^{s-1} (j+1) i_j = s_0}} \Pr \left\{ \frac{1}{N} \prod_{u=0}^{s-1} \left(\hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\mathbb{F}n}^u \hat{\boldsymbol{\delta}}_n \right)^{i_u} > \frac{\varepsilon}{\beta a_{\text{MAX}}^4 \phi_{\text{MAX}}^4 2^{s+1}} \right\}$$

Thanks to Property B on the convergence in probability, ζ_{2a} converges in probability with rate $o(N^{-2-\frac{4}{s}})$ at worst, i.e. $\forall \varepsilon > 0$,

$$\lim_{K=\beta N \rightarrow \infty} \Pr \left\{ \frac{\prod_{u=0}^{s-1} \hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\mathbb{F}n}^u \hat{\boldsymbol{\delta}}_n}{N} > \sqrt[s]{\frac{\varepsilon}{\beta 2^{s+1} a_{\text{MAX}}^4 \phi_{\text{MAX}}^4}} \right\} \leq o \left(\frac{1}{N^{2+\frac{4}{s}}} \right). \quad (55)$$

In fact, for $\varepsilon' = \frac{\varepsilon}{\beta 2^{s+1} a_{\text{MAX}}^4 \phi_{\text{MAX}}^4}$

$$\begin{aligned}
\Pr \left\{ \frac{\prod_{u=0}^{s-1} (\hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\text{F}n}^u \hat{\boldsymbol{\delta}}_n)^{i_u}}{N} > \varepsilon' \right\} &\leq \sum_{u=0}^{s-1} \Pr \left\{ \hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\text{F}n}^u \hat{\boldsymbol{\delta}}_n > \sqrt[s]{\varepsilon' N} \right\} \\
&\stackrel{(a)}{\leq} \sum_{u=0}^{s-1} \Pr \left\{ \left| \hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\text{F}n}^u \hat{\boldsymbol{\delta}}_n - \frac{\text{tr} \hat{\mathbf{R}}_{\text{F}n}^u}{N} \right| > \sqrt[s]{\varepsilon' N} - \frac{\text{tr} \hat{\mathbf{R}}_{\text{F}n}^u}{N} \right\} \\
&\stackrel{(b)}{\leq} \sum_{u=0}^{s-1} \frac{\mathbb{E} \left\{ \left| \hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{R}}_{\text{F}n}^u \hat{\boldsymbol{\delta}}_n - \frac{\text{tr} \hat{\mathbf{R}}_{\text{F}n}^u}{N} \right|^4 \right\}}{\sqrt[s]{(\varepsilon' N)^4}} \\
&\stackrel{(c)}{\leq} \frac{K_4 C^{(u)}}{N^2 ((N \varepsilon')^{\frac{1}{s}} - C^{(u)})^4}
\end{aligned} \tag{56}$$

where inequality (a) holds for N sufficiently large, inequality (b) follows from the Bienaymé inequality, and inequality (c) is a consequence of Lemma 9 in part I [1] and the bound on the eigenvalues moments of the matrix $\hat{\mathbf{R}}$.

Let us consider now the probability ζ_{2b} ,

$$\begin{aligned}
\zeta_{2b} &\leq \Pr \left\{ \frac{1}{N} \sum_{k=1}^K |(\hat{\mathbf{R}}^s)_{kk} - R_s(|a_k|^2, \tilde{\tau}_k)| > \frac{\varepsilon}{a_{\text{MAX}}^2 \phi_{\text{MAX}}^2} \right\} \\
&\leq \Pr \left\{ \max_k |(\hat{\mathbf{R}}^s)_{kk} - R_s(|a_k|^2, \tilde{\tau}_k)| > \frac{\varepsilon}{\beta a_{\text{MAX}}^2 \phi_{\text{MAX}}^2} \right\}
\end{aligned} \tag{57}$$

for $s = 1, \dots, \ell - 1$. Thanks to the assumption of the recursive step that $\forall \varepsilon' > 0$ and large $K = \beta N$, $\Pr\{|(\hat{\mathbf{R}}^s)_{kk} - R_s(|a_k|^2, \tilde{\tau}_k)| > \varepsilon'\} \leq o(N^{-2})$, $\zeta_{2b} \rightarrow o(N^{-2})$, i.e. it vanishes asymptotically as $N, K \rightarrow \infty$ with constant ratio with the same converge rate as $o(N^{-2})$ at worst. Therefore, (51) converges in probability with a rate as $o(N^{-2})$ for $N \rightarrow +\infty$, at worst. This convergence rate enables the application of the Borel-Cantelli lemma to prove that (51) converges almost surely.

The proof of the convergence (53) in probability one follows along similar lines.

Following the same approach as in the proof of Theorem 1 in [2], we can expand $(\hat{\mathbf{R}}^\ell)_{kk}$ and $\hat{\mathbf{T}}_{[nn]}^\ell$ as follows:

$$(\hat{\mathbf{R}}^\ell)_{kk} = \sum_{s=0}^{\ell-1} \hat{\mathbf{h}}_k^H \hat{\mathbf{T}}_{\sim k}^{\ell-s-1} \hat{\mathbf{h}}_k (\hat{\mathbf{R}}^s)_{kk} \quad \ell = 1, 2, \dots \tag{58}$$

$$\hat{\mathbf{T}}_{[nn]}^\ell = \sum_{s=0}^{\ell-1} \hat{\boldsymbol{\delta}}_n \hat{\mathbf{R}}_{\text{F}n}^{\ell-s-1} \hat{\boldsymbol{\delta}}_n^H \hat{\mathbf{T}}_{[nn]}^s \quad \ell = 1, 2, \dots \tag{59}$$

being $\hat{\mathbf{T}}^0$ and $\hat{\mathbf{R}}^0$ the identity matrices of dimensions $rN \times rN$ and $K \times K$, respectively.

Thanks to Property A and Property B of the convergence in probability of random sequences and the induction assumptions, the convergence in probability one of the sequences $\{(\hat{\mathbf{R}}^\ell)_{kk}\}$ and $\{\hat{\mathbf{T}}_{[nn]}^\ell\}$ reduces

to the following two steps. First we show the convergence in probability of $\hat{\mathbf{h}}_k^H \hat{\mathbf{T}}_{\sim k}^s \hat{\mathbf{h}}_k$ and $\hat{\boldsymbol{\delta}}_n^s \hat{\mathbf{R}}_{\neq n}^s \hat{\boldsymbol{\delta}}_n^H$ to a deterministic limit, respectively. Then, we show that the convergence holds with an appropriate convergence rate which enables the application of the Borel Cantelli lemma. Let us define

$$\zeta_3 = \hat{\mathbf{h}}_k^H \hat{\mathbf{T}}_{\sim k}^s \hat{\mathbf{h}}_k - \frac{|a_k|^2}{N} \text{tr} \Delta_{\phi,r}^H(\tilde{\tau}_k) \hat{\mathbf{T}}_{\sim k}^s \Delta_{\phi,r}(\tilde{\tau}_k).$$

Lemma 9 in part I [1] applied to the quadratic form $\hat{\mathbf{h}}_k^H \hat{\mathbf{T}}_{\sim k}^s \hat{\mathbf{h}}_k$ with $p = 4$ yields

$$\begin{aligned} \mathbb{E} |\zeta_3|^4 &< \frac{K_4 |a_k|^4}{N^3} \mathbb{E} \left(\text{tr}(\Delta_{\phi,r}^H(\tilde{\tau}_k) \hat{\mathbf{T}}_{\sim k}^s \Delta_{\phi,r}(\tilde{\tau}_k))^4 \right) \\ &\leq \frac{K_4}{N^3} a_{\text{MAX}}^8 \phi_{\text{MAX}}^8 \text{tr}(\hat{\mathbf{T}}_{\sim k}^{4s}). \end{aligned} \quad (60)$$

Thanks to the bound on the eigenvalues moments of the matrix $\hat{\mathbf{T}}$, $\lim_{K=\beta N \rightarrow \infty} \frac{1}{N} \mathbb{E}(\text{tr} \hat{\mathbf{T}}_{\sim k}^{4s})$ is almost sure upper bounded $\forall s$ as $N = \beta K \rightarrow +\infty$. Therefore, $\mathbb{E} |\zeta_3|^4 \rightarrow 0$ as $K, N \rightarrow \infty$ with $\frac{K}{N} \rightarrow \beta$ and $\hat{\mathbf{h}}_k^H \hat{\mathbf{T}}_{\sim k}^s \hat{\mathbf{h}}_k$ converges in mean square sense, and thus in probability. Furthermore, the Bienaymé inequality implies that $\Pr\{|\zeta_3| > \varepsilon\} \leq o(N^{-2})$ as $N \rightarrow +\infty$. Thanks to (53)

$$\begin{aligned} \lim_{N=\beta K \rightarrow \infty} \frac{|a_k|^2}{N} \text{tr} \Delta_{\phi,r}^H(\tilde{\tau}_k) \hat{\mathbf{T}}_{\sim k}^s \Delta_{\phi,r}(\tilde{\tau}_k) &= \frac{\lambda}{2\pi} \int_0^{2\pi} \Delta_{\phi,r}^H(\Omega, \tau) \mathbf{T}_s(\Omega) \Delta_{\phi,r}(\Omega, \tau) d\Omega \Big|_{(\lambda, \tau)=(|a_k|^2, \tilde{\tau}_k)} + o(N^{-2}) \\ &= g(\mathbf{T}_s, \lambda, \tau) + o(N^{-2}). \end{aligned} \quad (61)$$

then

$$\Pr\{|\hat{\mathbf{h}}_k^H \hat{\mathbf{T}}_{\sim k}^s \hat{\mathbf{h}}_k - g(\mathbf{T}_s, \lambda, \tau)| > \varepsilon\} \rightarrow o(N^{-2}) \quad (62)$$

for property A. Thanks to the convergence rate in (62) and the Borel Cantelli lemma, the almost sure convergence (52) follows.

The convergence in probability one of the diagonal blocks $\hat{\mathbf{T}}_{[nn]}^\ell$ can be proven in a similar way. More specifically, it can be shown that the $r \times r$ block $\hat{\boldsymbol{\delta}}_n^s \hat{\mathbf{R}}_{\neq n}^s \hat{\boldsymbol{\delta}}_n^H$ converges to the $r \times r$ deterministic matrix

$$\mathbf{f}(R_s, \Omega) = \beta \int \lambda \Delta_{\phi,r}(\Omega, \tau) \Delta_{\phi,r}^H(\Omega, \tau) R_s(\lambda, \tau) dF_{|\mathbf{A}|^2, T}(\lambda, \tau). \quad (63)$$

such that $\Pr\left\{\left|(\hat{\boldsymbol{\delta}}_n)_u \hat{\mathbf{R}}_{\neq n}^s (\hat{\boldsymbol{\delta}}_n)_v - (\mathbf{f}(R_s, \Omega))_{u,v}\right| > \varepsilon\right\} \rightarrow o(N^{-2})$.

Finally, by making use of equations (58) and (59) and the definitions (52), (54), (63), and (61) we obtain

$$R_\ell(\lambda, \tau) = \sum_{s=0}^{\ell-1} g(\mathbf{T}_{\ell-s-1}, \lambda, \tau) R_s(\lambda, \tau) \quad \ell = 1, 2, \dots \quad (64)$$

and

$$\mathbf{T}_\ell(\Omega) = \sum_{s=0}^{\ell-1} \mathbf{f}(R_{\ell-s-1}, \Omega) \mathbf{T}_s(\Omega) \quad \ell = 1, 2, \dots \quad (65)$$

with $g(\mathbf{T}_s, \lambda, \tau)$ and $\mathbf{f}(R_s, \Omega)$ given in (61) and (63), respectively. Consistently to the definitions of $\hat{\mathbf{T}}^0$ and $\hat{\mathbf{R}}^0$, $\mathbf{T}_0(\Omega) = \mathbf{I}_r$, being \mathbf{I}_r the $r \times r$ identity matrix and $R_0(\lambda) = 1$.

Then, $g(R_0, \lambda, \tau) = \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \Delta_{\phi,r}^H(\Omega, \tau) \Delta_{\phi,r}(\Omega, \tau) d\Omega$ and $\mathbf{f}(\mathbf{T}_0, \Omega) = \beta \int \lambda \Delta_{\phi,r}(\Omega, \tau) \Delta_{\phi,r}^H(\Omega, \tau) dF_{|\mathbf{A}|^2, T}(\lambda, \tau)$ and (64) and (65) reduce to the asymptotic limits $R_1(\lambda, \tau)$ and $\mathbf{T}_1(\Omega)$ already derived in *step 1*. Therefore, we can begin the recursion with $\ell = 0$, $R_0(\lambda, \tau) = 1$ and $\mathbf{T}_0(\Omega) = \mathbf{I}_r$.

Properties A, B, and C, the induction assumptions, relations (58) and (64), the convergence rates $\zeta_2 \rightarrow o(N^{-2})$, $\Pr\{\zeta_3 > \varepsilon\} \leq o(N^{-2})$, and the Borel Cantelli lemma yield (44). The proof of (45) follows immediately along similar lines.

This concludes the proof of Theorem 1.

APPENDIX II

PROOF OF COROLLARY 1

Corollary 1 is derived by specializing Theorem 1 to a unitary Fourier transform $\Phi(\omega)$ with bandwidth $B \leq \frac{r}{2T_c}$. Let us recall here that the unitary Fourier transform in the discrete time domain is given by

$$\phi(\Omega, \tau) = \frac{1}{T_c} e^{j\frac{\tau}{T_c}\Omega} \sum_{s=-\text{sign}(\Omega)\lfloor \frac{r-1}{2} \rfloor}^{\text{sign}(\Omega)\lfloor \frac{r}{2} \rfloor} e^{j2\pi\frac{\tau}{T_c}s} \Phi^*\left(\frac{\Omega + 2\pi s}{T_c}\right) \quad \text{for } |\Omega| \leq \pi. \quad (66)$$

The matrix $\mathbf{Q}(\Omega, \tau) = \Delta_{\phi,r}(\Omega, \tau) \Delta_{\phi,r}^H(\Omega, \tau)$, with $\Delta_{\phi,r}(\Omega, \tau)$ defined in (26), can be decomposed as $\mathbf{Q}(\Omega, \tau) = \mathbf{Q}(\Omega) + \overline{\mathbf{Q}}(\Omega, \tau)$ with the elements of $\mathbf{Q}(\Omega)$ and $\overline{\mathbf{Q}}(\Omega, \tau)$ defined by

$$(\mathbf{Q}(\Omega))_{k,\ell} = \frac{1}{T_c^2} \sum_{s=-\text{sign}(\Omega)\lfloor \frac{r-1}{2} \rfloor}^{\text{sign}(\Omega)\lfloor \frac{r}{2} \rfloor} \left| \Phi\left(\frac{\Omega + 2\pi s}{T_c}\right) \right|^2 e^{-j\frac{k-\ell}{r}(\Omega + 2\pi s)} \quad \text{for } |\Omega| \leq \pi, \quad (67)$$

and

$$(\overline{\mathbf{Q}}(\Omega, \tau))_{k,\ell} = \frac{1}{T_c^2} \sum_{\substack{s,u=-\text{sign}(\Omega)\lfloor \frac{r-1}{2} \rfloor \\ s \neq u}}^{\text{sign}(\Omega)\lfloor \frac{r}{2} \rfloor} \Phi\left(\frac{\Omega + 2\pi u}{T_c}\right) \Phi^*\left(\frac{\Omega + 2\pi s}{T_c}\right) e^{-j2\pi\frac{\tau}{T_c}(s-u)} e^{-j\left(\frac{k-1}{r}(\Omega - 2\pi s) - \frac{\ell-1}{r}(\Omega - 2\pi u)\right)} \quad \text{for } |\Omega| \leq \pi, \quad (68)$$

respectively.

Equations (24) and (25) can be rewritten as

$$\begin{aligned} f(R_s, \Omega) &= \beta \mathbf{Q}(\Omega) \int \lambda R_s(\lambda, \tau) dF_{|\mathbf{A}|^2, T}(\lambda, \tau) \\ &\quad + \beta \int \lambda R_s(\lambda, \tau) \overline{\mathbf{Q}}(\Omega, \tau) dF_{|\mathbf{A}|^2, T}(\lambda, \tau), \end{aligned} \quad -\pi \leq \Omega \leq \pi \quad (69)$$

$$g(\mathbf{T}_s, \lambda, \tau) = \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \text{tr}(\mathbf{T}_s(\Omega) \mathbf{Q}(\Omega)) d\Omega + \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \text{tr}(\mathbf{T}_s(\Omega) \overline{\mathbf{Q}}(\Omega, \tau)) d\Omega, \quad (70)$$

respectively. If the conditions of Corollary 1 are satisfied, i.e. if $B \leq \frac{r}{2T_c}$ and τ is uniformly distributed in $[0, T_c]$, it can be shown that

- $R_\ell(\lambda, \tau)$, $\ell \in \mathbb{Z}^+$, are independent of τ and
- $\mathbf{T}_\ell(\Omega)$ is a matrix of the form (71).

$$\mathbf{B} = \mathbf{B}(\Omega) = \begin{bmatrix} b_0 & b_1 e^{j\frac{\Omega}{r}} & \dots & \dots & b_{r-1} e^{j\frac{(r-1)}{r}\Omega} \\ b_{r-1} e^{-j\frac{\Omega}{r}} & b_0 & b_1 e^{j\frac{\Omega}{r}} & \dots & b_{r-2} e^{j\frac{(r-2)}{r}\Omega} \\ \dots & \ddots & \ddots & \ddots & \ddots \\ b_1 e^{-j\frac{(r-1)}{r}\Omega} & \ddots & \ddots & b_{r-1} e^{-j\frac{\Omega}{r}} & b_0 \end{bmatrix}, \quad (71)$$

being $b_0 = b_0(\Omega)$, $b_1 = b_1(\Omega)$, \dots , $b_{r-1} = b_{r-1}(\Omega)$, eventually functions of Ω .

These properties can be proven by strong induction. It is straightforward to verify that they are satisfied for $s = 0$. In fact, $R_0(\lambda, \tau) = 1$ is independent of τ and $\mathbf{T}_0(\Omega) = \mathbf{I}$ is of the form (71) with $b_0 = 1$ and $b_i(\Omega) = 0$ with $i = 1, \dots, r-1$. By appealing to Lemma 1 in part I [1] Appendix I $\text{tr}(\overline{\mathbf{Q}}(\Omega, \tau)) = 0$ and $g(\mathbf{T}_0, \lambda, \tau) = \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \text{tr}(\mathbf{Q}(\Omega)) d\Omega$. Hence, $g(\mathbf{T}_0, \lambda, \tau)$ is independent of τ .

The induction step is proven using the following induction assumptions:

- For $s = 0, 1, \dots, \ell-1$, $R_s(\lambda, \tau)$ is independent of τ ;
- For $s = 0, 1, \dots, \ell-1$, $\mathbf{T}_s(\Omega)$ is of the form (71).

Thanks to the form (71) of $\mathbf{T}_s(\Omega)$, $s = 1, \dots, \ell-1$, given by the induction assumptions and by applying Lemma I in part I Appendix I we have $\text{tr}(\mathbf{T}_s(\Omega) \overline{\mathbf{Q}}(\Omega, \tau)) = 0$, for $s = 0, 1, \dots, \ell-1$. Then, (70) reduces to $g(\mathbf{T}_s, \lambda, \tau) = \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \text{tr}(\mathbf{T}_s(\Omega) \mathbf{Q}(\Omega)) d\Omega$ and $g(\mathbf{T}_s, \lambda, \tau)$ is independent of τ for $s = 0, 1, \dots, \ell-1$. Therefore, all quantities that appear in the right hand side of (22) are independent of τ and $R_\ell(\lambda, \tau)$ is also independent of τ . In the following we will shortly write $R_\ell(\lambda)$ and $g(\mathbf{T}_s, \lambda)$ instead of $R_\ell(\lambda, \tau)$ and $g(\mathbf{T}_s, \lambda, \tau)$. Thanks to the fact that (i) $R_s(\lambda, \tau)$ is independent of τ and (ii) λ and τ are statistically independent with τ uniformly distributed, (69) can be rewritten as

$$f(R_s, \Omega) = \beta \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2} \left(\mathbf{Q}(\Omega) + \frac{1}{T_c} \int_0^{T_c} \overline{\mathbf{Q}}(\Omega, \tau) d\tau \right). \quad (72)$$

It is straightforward to verify that $\int_0^{T_c} \overline{\mathbf{Q}}(\Omega, \tau) d\tau = 0$ from the definition of $\overline{\mathbf{Q}}(\Omega, \tau)$ in (68). Then,

$$\begin{aligned} \mathbf{f}(R_s, \Omega) &= \beta \mathbf{Q}(\Omega) \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2}(\lambda) \\ &= f(R_s) \mathbf{Q}(\Omega) \end{aligned} \quad (73)$$

with $f(R_s) = \beta \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2}(\lambda)$. Substituting (73) in (23) yields

$$\mathbf{T}_\ell(\Omega) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \mathbf{Q}(\Omega) \mathbf{T}_s(\Omega), \quad -\pi \leq \Omega \leq \pi. \quad (74)$$

Since $\mathbf{T}_s(\Omega)$ is of form (71), the conditions of Lemma 2 in part I Appendix I are satisfied for $\mathbf{B} = \mathbf{T}_s(\Omega)$. This implies that $\mathbf{Q}(\Omega) \mathbf{T}_s(\Omega)$ is also of the form (71). Since $\mathbf{T}_\ell(\Omega)$ is a linear combination of matrices of the form (71), $\mathbf{T}_\ell(\Omega)$ is also a matrix of the form (71). Then, the statement of the strong induction is proven.

Thanks to the properties shown by strong induction, the recursive equations in Theorem (1) reduce to the following set of recursive equations:

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} g(\mathbf{T}_{\ell-s-1}, \lambda) R_s(\lambda) \quad (75)$$

$$\mathbf{T}_\ell(\Omega) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \mathbf{Q}(\Omega) \mathbf{T}_s(\Omega) \quad -\pi \leq \Omega \leq \pi \quad (76)$$

$$f(R_s) = \beta \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2}(\lambda), \quad (77)$$

$$g(\mathbf{T}_s, \lambda) = \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \text{tr}(\mathbf{T}_s(\Omega) \mathbf{Q}(\Omega)) d\Omega \quad (78)$$

with $\mathbf{T}_0(\Omega) = \mathbf{I}_r$ and $R_0(\lambda) = 1$.

Then, applying again Theorem 1 we obtain the following convergence in probability one

$$\lim_{K=\beta N \rightarrow \infty} (\hat{\mathbf{R}}^\ell)_{kk} = R_\ell(\lambda)|_{\lambda=|a_k|^2}.$$

From (76) and $\mathbf{T}_0(\Omega) = \mathbf{I}_r$ it is apparent that $\mathbf{T}_\ell(\Omega)$ is a polynomial in $\mathbf{Q}^s(\Omega)$, for $s = 0, 1, \dots, \ell$. Then, $\mathbf{T}_\ell(\Omega)$ has the same eigenvectors as $\mathbf{Q}(\Omega)$ and it can be written as $\mathbf{T}_\ell(\Omega) = \mathbf{U}(\Omega) \mathbf{\Lambda}_\ell(\Omega) \mathbf{U}^H(\Omega)$ where $\mathbf{\Lambda}_\ell(\Omega)$ is a diagonal matrix with diagonal elements $t_{\ell,1}, t_{\ell,2}, \dots, t_{\ell,r}$ and

$$\mathbf{U}(\Omega) = \left(e \left(\Omega - \text{sign}(\Omega) 2\pi \left\lfloor \frac{r-1}{2} \right\rfloor \right), \dots, e(\Omega) \dots e \left(\Omega + \text{sign}(\Omega) 2\pi \left\lfloor \frac{r}{2} \right\rfloor \right) \right) \quad (79)$$

with $e(\Omega)$ r -dimensional column vector defined by

$$e(\Omega) = \frac{1}{\sqrt{r}} \left(1, e^{-j\frac{\Omega}{r}}, \dots, e^{-j\frac{r-1}{r}\Omega} \right)^T.$$

By making use of the eigenvalue decomposition of the matrix $\mathbf{Q}(\Omega)$ in part I Appendix I Lemma 3 the matrix equation (76) reduces to r scalar equations

$$t_{\ell,u}(\Omega) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \frac{r}{T_c^2} \left| \Phi \left(\frac{\Omega}{T_c} - \text{sign}(\Omega) \frac{2\pi}{T_c} \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \right) \right|^2 t_{s,u}(\Omega) \quad u = 1, \dots, r \quad \text{and} \quad |\Omega| \leq \pi.$$

By substituting $y = \Omega - \text{sign}(\Omega) 2\pi \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right)$ for $|\Omega| \leq \pi$ we obtain

$$t_{\ell,u} \left(y + 2\pi \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \right) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \frac{r}{T_c^2} \left| \Phi \left(\frac{y}{T_c} \right) \right|^2 t_{s,u} \left(y + 2\pi \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \right) \quad (80)$$

for $0 \leq y + 2\pi \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \leq \pi$ and

$$t_{\ell,u} \left(y - 2\pi \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \right) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \frac{r}{T_c^2} \left| \Phi \left(\frac{y}{T_c} \right) \right|^2 t_{s,u} \left(y - 2\pi \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \right) \quad (81)$$

for $-\pi \leq y - 2\pi \left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \leq 0$. Then, for $u = 1, \dots, r$, the r functions (80) and (81) defined in not overlapping intervals in $[-2\pi r, 2\pi r]$ can be combined in a unique scalar functions $T'_\ell(y)$ in the interval $|y| \leq 2\pi r$ satisfying the recursive equation

$$T'_\ell(y) = \sum_{s=0}^{\ell-1} \frac{r}{T_c^2} f(R_{\ell-s-1}) \left| \Phi \left(\frac{y}{T_c} \right) \right|^2 T'_s(y).$$

Similar arguments applied to (78) yield

$$g(T_s, \lambda) = \frac{\lambda}{2\pi} \int_{-r\pi}^{r\pi} \frac{r}{T_c^2} T'_s(y) \left| \Phi \left(\frac{y}{T_c} \right) \right|^2 dy.$$

The substitutions $\omega = \frac{y}{T_c}$ and $T'_\ell(\omega T_c) = T'_\ell(\omega)$ yield to the recursive equations in Corollary 1.

This concludes the derivation of Corollary 1 from Theorem 1.

APPENDIX III

DERIVATION OF ALGORITHM 1

Algorithm 1 can be derived from the recursive equations of Corollary 1 by using the following substitutions⁶:

λ	\rightarrow	z
$R_s(\lambda)$	\rightarrow	$\rho_s(z)$
$\lambda R_s(\lambda)$	\rightarrow	$v_s(z)$
$E(\lambda R_s(\lambda)) = \frac{1}{\beta} f(R_s)$	\rightarrow	V_s
$\frac{1}{T_c} \Phi(\omega) ^2$	\rightarrow	y
$T_s(\cdot)$	\rightarrow	$\mu_s(y)$
$\frac{r}{T_c} \Phi(\omega) ^2 T_s(\omega)$	\rightarrow	$u_s(y)$
$\frac{r}{2\pi T_c} \int_{-2\pi B}^{2\pi B} \Phi(\omega) ^2 T_s(\omega) d\omega$	\rightarrow	U_s

Then, the initial step is obtained by defining $\mu_0(y) = 1$ and $\rho_0(z) = 1$. The recursive equations in step ℓ are obtained by using the previous substitutions. In order to derive U_s let us observe that $\frac{1}{T_c} |\Phi(\omega)|^2 T_s(\omega)$ is a polynomial in $y = \frac{1}{T_c} |\Phi(\omega)|^2$ of degree $s + 1$. Then, U_s is a linear combination of $\frac{\mathcal{E}_n}{T_c}$ where

$$\mathcal{E}_n = \frac{1}{2\pi T_c^{n-1}} \int_{-2\pi B}^{2\pi B} |\Phi(\omega)|^{2n} d\omega$$

The coefficients of the linear combination are obtained by expanding $u_s(y)$ as a polynomial in y .

We conclude the derivation of Algorithm 1 by summarizing the previous considerations and substitutions:

•

$$\begin{aligned} \rho_\ell(z) &= \sum_{s=0}^{\ell-1} z U_{\ell-s-1} \rho_s(z) \\ \mu_\ell(y) &= \frac{r}{T_c} \sum_{s=0}^{\ell-1} \beta y V_{\ell-s-1} \mu_s(y). \end{aligned}$$

- U_s and V_s are obtained from $u_s(y) = y \mu_s(y)$ and $v_s(z) = z \rho_s(z)$, respectively by
 - expanding $u_s(y)$ and $v_s(z)$ as polynomials in y and z , respectively,

⁶Note that the substitution of λ with z is redundant. It is used to obtain polynomials in the commonly used variable z .

– replacing the monomials y^n and z^n , $n \in \mathbb{Z}^+$ with $\frac{\varepsilon_n}{T_c}$ and $m_{|\mathbf{A}|^2}^{(s)}$, respectively.

Then, $R_\ell(\lambda) = \rho_\ell(\lambda)$ and the eigenvalue moment $m_{\mathcal{R}}^{(\ell)} = \mathbb{E}\{R_\ell(\lambda)\}$ is obtained by replacing all monomials z, z^2, \dots, z^ℓ in the polynomial $\rho_\ell(z)$ by the moments $m_{|\mathbf{A}|^2}^1, m_{|\mathbf{A}|^2}^2, \dots, m_{|\mathbf{A}|^2}^\ell$, respectively.

APPENDIX IV

PROOF OF THEOREM 2

The proof of Theorem 2 follows along the line of the proof of Theorem 1. As in the proof of Theorem 1 we can focus on the spreading matrix $\bar{\mathbf{S}}$ in (39) and the autocorrelation $\bar{\mathbf{R}}$.

For a signal with bandwidth $B \leq \frac{1}{2T_c}$,

$$\underline{\phi}(\Omega, \tau) = \frac{1}{T_c} e^{j\frac{\tau\Omega}{T_c}} \Phi^* \left(\frac{\Omega}{T_c} \right) \quad |\Omega| \leq \pi$$

and $\phi(\Omega, \tau) = \underline{\phi}(\Omega - 2\pi \lfloor \frac{\Omega}{\pi} \rfloor, \tau)$ for any Ω . Correspondingly, we define

$$\underline{\Delta}_{\phi,r}(\Omega, \tau) = \frac{1}{T_c} \Phi \left(\frac{\Omega}{T_c} \right) e^{-j\frac{\tau\Omega}{T_c}} \mathbf{e}(\Omega), \quad |\Omega| \leq \pi$$

with $\mathbf{e}(\Omega) = (1, e^{j\frac{\Omega}{r}}, \dots, e^{j\frac{(r-1)}{r}\Omega})$ and

$$\Delta_{\phi,r}(\Omega, \tau) = \underline{\Delta}_{\phi,r}(\Omega - 2\pi \lfloor \frac{\Omega}{\pi} \rfloor, \tau) \quad \text{for any } \Omega.$$

We adopt here the same notation as in the proof of Theorem 1. Then, the $K \times K$ diagonal matrix ∇_{nt} , for $t = 1, \dots, r$ and $n = 1, \dots, N$ is given by

$$\nabla_{nt} = \frac{1}{T_c} \Phi^* \left(\frac{j2\pi}{T_c} \underline{n} \right) e^{-\frac{j2\pi \underline{n}(t-1)}{r}} \text{diag} \left(e^{\frac{j2\pi \underline{n} \tilde{\tau}_1}{T_c}}, e^{\frac{j2\pi \underline{n} \tilde{\tau}_2}{T_c}}, \dots, e^{\frac{j2\pi \underline{n} \tilde{\tau}_K}{T_c}} \right)$$

with $\underline{n} = \frac{n-1}{N} - \lfloor 2\frac{n-1}{N} \rfloor$ and $\Delta_{\phi,r}(\tilde{\tau}_k)$ is the $rN \times N$ block diagonal matrix with n diagonal block $\Delta_{\phi,r}(\underline{n}, \tilde{\tau}_k)$.

We develop the proof by strong induction as in Theorem 1 with similar initial step and similar induction step.

Step 1: In this case

$$\hat{\mathbf{R}}_{kk} = |a_k|^2 \mathbf{s}_k^H \Delta_{\phi,r}^H(\tilde{\tau}_k) \Delta_{\phi,r}(\tilde{\tau}_k) \mathbf{s}_k = |a_k|^2 \mathbf{s}_k^H \Phi \mathbf{s}_k$$

where Φ is a matrix independent of $\tilde{\tau}_k$ and the n^{th} element is given by $\Phi_{nn} = \frac{r}{T_c} \left| \Phi \left(\frac{j2\pi \underline{n}}{T_c} \right) \right|^2$.

By following the same approach as in Theorem 1 it results $\forall \varepsilon > 0$

$$\Pr \left\{ \left| \hat{\mathbf{R}}_{kk} - \frac{r|a_k|^2}{T_c N} \sum_{n=0}^{N-1} \left| \Phi \left(\frac{j2\pi \underline{n}}{T_c} \right) \right|^2 \right| > \varepsilon \right\} \leq \frac{K_4 |a_k|^4 \Delta_{\text{MAX}}^4}{N^2 \varepsilon^4}$$

being $\Delta_{\text{MAX}} = \max_{\Omega \in [-\pi, \pi]} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2$ and

$$\begin{aligned} R_1(\lambda)|_{\lambda=|a_k|^2} &= \lim_{K=\beta N \rightarrow \infty} \frac{|a_k|^2}{N} \sum_{\ell=0}^{N-1} \left| \Phi \left(\frac{2\pi}{T_c} \left(\frac{n}{N} - \left\lfloor \frac{2n}{N} \right\rfloor \right) \right) \right|^2 \\ &= \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 d\Omega \Big|_{\lambda=|a_k|^2}. \end{aligned} \quad (82)$$

Furthermore, $\Pr \left\{ |\hat{\mathbf{R}}_{kk} - R_1(|a_k|^2)| > \varepsilon \right\} \leq o(N^{-2})$ with consequent convergence in probability one by the Borel Cantelli lemma.

Similarly, $(\hat{\mathbf{T}}_{[nn]})_{uv}$, the (u, v) -element of the matrix $\hat{\mathbf{T}}_{[nn]}$ is given by

$$\begin{aligned} \hat{\mathbf{T}}_{[nn]} &= \hat{\boldsymbol{\sigma}}_n \mathbf{A} \nabla_{n,u} \nabla_{n,v}^H \mathbf{A}^H \hat{\boldsymbol{\sigma}}_n^H \\ &= \frac{1}{T_c} \left| \Phi \left(\frac{2\pi \underline{n}}{T_c} \right) \right|^2 e^{-j2\pi \underline{n} \frac{v-u}{r}} \hat{\boldsymbol{\sigma}}_n \mathbf{A} \mathbf{A}^H \hat{\boldsymbol{\sigma}}_n^H. \end{aligned} \quad (83)$$

As in Theorem 1 it can be shown that

$$\Pr \left\{ \left| (\hat{\mathbf{T}}_{[nn]})_{uv} - \frac{1}{NT_c} \left| \Phi \left(\frac{2\pi \underline{n}}{T_c} \right) \right|^2 e^{-j2\pi \underline{n} \frac{v-u}{r}} \text{tr}(\mathbf{A} \mathbf{A}^H) \right| > \varepsilon \right\} \leq \frac{K_4 T_{\text{MAX}}^4}{N^2 \varepsilon^4}$$

with $T_{\text{MAX}} = \left(\max_{\Omega \in [-\pi, \pi]} \left| \Phi \left(\frac{2\pi \underline{n}}{T_c} \right) \right|^2 \right) (\sup_K \max_k |a_k|^2)$ and the following convergence in probability holds

$$\begin{aligned} \lim_{K=\beta N \rightarrow \infty} (\hat{\mathbf{T}}_{[nn]})_{uv} &= \lim_{K=\beta N \rightarrow \infty} \frac{\beta}{T_c K} \left| \Phi \left(\frac{2\pi \underline{n}}{T_c} \right) \right|^2 e^{-j2\pi \underline{n} \frac{v-u}{r}} \sum_{k=1}^K |a_k|^2 \\ &= \frac{\beta}{T_c} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 e^{-j2\pi \underline{n} \frac{v-u}{r}} \int \lambda dF_{|\mathbf{A}|^2}(\lambda) \end{aligned}$$

with $\Omega = 2\pi \lim_{N \rightarrow \infty} \underline{n}$ and $|\Omega| \leq \pi$. Thus, the diagonal block converges in probability as follows

$$\begin{aligned} \mathbf{T}_1(\Omega) &= \lim_{K=\beta N \rightarrow \infty} (\hat{\mathbf{T}}_{[nn]})_{uv} \\ &= \frac{\beta}{T_c} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 \int \lambda dF_{|\mathbf{A}|^2}(\lambda) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) \end{aligned} \quad (84)$$

Furthermore,

$$\Pr \left\{ \left| (\hat{\mathbf{T}}_{[nn]})_{uv} - (\mathbf{T}_1(\Omega))_{uv} \right| > \varepsilon \right\} \leq o(N^{-2})$$

with consequent convergence in probability one by the Borel Cantelli lemma. This concludes the first step of the induction.

Step ℓ : Let us observe that

$$\begin{aligned}\vartheta_1 &= \frac{1}{N} \text{tr} \mathbf{A} \nabla_{n,u} \widehat{\mathbf{R}}_{\mathbb{F}n}^s \nabla_{n,u}^H \mathbf{A}^H \\ &= \frac{e^{-j2\pi \underline{n} \frac{u-v}{r}}}{N} \sum_{k=1}^K \frac{|a_k|^2}{T_c^2} \left| \Phi \left(\frac{2\pi \underline{n}}{T_c} \right) \right|^2 (\widehat{\mathbf{R}}_{\mathbb{F}n}^s)_{kk}\end{aligned}$$

and

$$\begin{aligned}\vartheta_2 &= \frac{|a_k|^2}{N} \text{tr} \Delta_{\Phi,r}^H(\tilde{\tau}_k) \widehat{\mathbf{T}}_{\sim k}^s \Delta_{\Phi,r}(\tilde{\tau}_k) \\ &= \frac{|a_k|^2}{N} \sum_{n=1}^N \frac{1}{T_c^2} \left| \Phi \left(\frac{2\pi \underline{n}}{T_c} \right) \right|^2 e^H(2\pi \underline{n}) (\widehat{\mathbf{T}}_{\sim k}^s)_{nn} e(2\pi \underline{n}).\end{aligned}$$

By following the same approach as in Theorem 1 it can be shown that ϑ_1 and ϑ_2 converge almost surely to the following limits

$$\lim_{K=\beta N \rightarrow \infty} \vartheta_1 = \frac{\beta}{T_c^2} e^{-j2\pi \underline{n} \frac{u-v}{r}} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2}(\lambda)$$

and

$$\lim_{K=\beta N \rightarrow \infty} \vartheta_2 = \frac{\lambda}{2\pi T_c^2} \int_{-\pi}^{\pi} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 e^H(\Omega) \mathbf{T}_s(\Omega) e(\Omega) d\Omega \Big|_{\lambda=|a_k|^2}$$

with $R_s(\lambda)|_{\lambda=|a_k|^2} = \lim_{K=\beta N \rightarrow \infty} (\widehat{\mathbf{R}}^s)_{kk}$ and $\mathbf{T}_s(\Omega)| = \lim_{K=\beta N \rightarrow \infty} \widehat{\mathbf{T}}_{[nn]}^s$ given by the recursion assumptions.

Additionally, it can be shown that the following almost sure convergence holds

$$\begin{aligned}g(\mathbf{T}_s, \lambda)|_{\lambda=|a_k|^2} &= \lim_{K=\beta N \rightarrow \infty} \widehat{\mathbf{h}}_k^H \widehat{\mathbf{T}}_{\sim k}^s \widehat{\mathbf{h}}_k \\ &= \frac{\lambda}{2\pi T_c} \int_{-\pi}^{\pi} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 e^H(\Omega) \mathbf{T}_s(\Omega) e(\Omega) d\Omega \Big|_{\lambda=|a_k|^2}\end{aligned}\tag{85}$$

and

$$\begin{aligned}\mathbf{f}(R_s, \Omega) &= \lim_{K=\beta N \rightarrow \infty} \widehat{\boldsymbol{\delta}}_n^H \widehat{\mathbf{R}}_{\mathbb{F}n}^s \widehat{\boldsymbol{\delta}}_n^H \\ &= \frac{\beta}{T_c^2} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 e(\Omega) e^H(\Omega) \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2}(\lambda)\end{aligned}\tag{86}$$

Additionally, the convergence satisfies the bounds

$$\Pr \left\{ |\widehat{\mathbf{h}}_k^H \widehat{\mathbf{T}}_{\sim k}^s \widehat{\mathbf{h}}_k - g(\mathbf{T}_s, |a_k|^2)| > \varepsilon \right\} < o(N^{-2})$$

and

$$\Pr \left\{ |(\widehat{\boldsymbol{\delta}}_n)_u \widehat{\mathbf{R}}_{\mathbb{F}n}^s (\widehat{\boldsymbol{\delta}}_n^H)_v - (\mathbf{f}(R_s, \Omega))_{u,v}| > \varepsilon \right\} < o(N^{-2})$$

for large N and $\forall \varepsilon$.

The recursion assumptions and the limits (85) and (86) in (58) and (59) yield

$$\begin{aligned} R_\ell(\lambda)|_{\lambda=|a_k|^2} &= \sum_{s=0}^{\ell-1} g(\mathbf{T}_{\ell-s-1}, \lambda) R_s(\lambda) \\ &= \sum_{s=0}^{\ell-1} R_s(\lambda) \frac{\lambda}{2\pi T_c^2} \int_{-\pi}^{\pi} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 \text{tr} (\mathbf{T}_s(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega)) d\Omega \Big|_{\lambda=|a_k|^2} \end{aligned} \quad (87)$$

and

$$\begin{aligned} \mathbf{T}_\ell(\Omega) &= \sum_{s=0}^{\ell-1} \mathbf{f}(R_{\ell-s-1}, \Omega) \mathbf{T}_s(\Omega) \\ &\quad \sum_{s=0}^{\ell-1} \frac{\beta}{T_c^2} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2}(\lambda) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) \mathbf{T}_s(\Omega) \end{aligned} \quad (88)$$

where $R_0(\lambda) = 1$ and $\mathbf{T}_0(\Omega) = \mathbf{I}_r$. With a similar approach as in Theorem 1 it can be proven that for large N and $\forall \varepsilon > 0$

$$\Pr \left\{ \left| \widehat{\mathbf{R}}_{kk}^\ell - R_\ell(|a_k|^2) \right| > \varepsilon \right\} \leq o(N^{-2})$$

and

$$\Pr \left\{ \left| (\widehat{\mathbf{T}}_{[nn]}^\ell)_{uv} - (\mathbf{T}_\ell(\Omega))_{uv} \right| > \varepsilon \right\} \leq o(N^{-2}).$$

In contrast to Theorem 1 the recursive equations (87), (88), (85), and (86) are independent of the time delay $\widetilde{\tau}_k$.

The recursive equations can be further simplified by observing that $(\mathbf{e}(\Omega) \mathbf{e}^H(\Omega))^m = r^{m-1} \mathbf{e}(\Omega) \mathbf{e}^H(\Omega)$. Then, it is straightforward to verify by recursion that the matrix $\mathbf{T}_s(\Omega)$, $s = 1, 2, \dots, \ell - 1$, is proportional to the matrix $\mathbf{e}(\Omega) \mathbf{e}^H(\Omega)$ and we can express it as $\mathbf{T}_s(\Omega) = T_s(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega)$, $s = 1, 2, \dots$. Thus, the

recursive equations can be rewritten as

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} g(\mathbf{T}_{\ell-s-1}, \lambda) R_s(\lambda)$$

$$T_\ell(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) = \sum_{s=1}^{\ell-1} \mathbf{f}(R_{\ell-s-1}, \Omega) T_s(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) + \mathbf{f}(R_{\ell-1}, \Omega) \mathbf{T}_0(\Omega) \quad \ell = 1, 2, \dots \quad (89)$$

$$\mathbf{f}(R_s, \Omega) = f(R_s, \Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) \quad (90)$$

$$f(R_s, \Omega) = \frac{\beta}{T_c^2} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 \int \lambda R_s(\lambda) d F_{|\mathbf{A}|^2}(\lambda) \quad -\pi \leq \Omega \leq \pi$$

$$g(T_s, \lambda) = \begin{cases} \frac{r^2 \lambda}{2\pi T_c^2} \int_{-\pi}^{\pi} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 T^s(\Omega) d\Omega & s = 1, 2, \dots \\ \frac{r\lambda}{2\pi T_c^2} \int_{-\pi}^{\pi} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 d\Omega & s = 0. \end{cases}$$

with $\mathbf{T}_0(\Omega) = \mathbf{I}_r$ and $R_0(\lambda) = 1$.

Substituting (90) in (89) we obtain

$$T_\ell(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) = \sum_{s=1}^{\ell-1} f(R_{\ell-s-1}, \Omega) T_s(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) + f(R_{\ell-1}, \Omega) \mathbf{T}_0(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega)$$

$$= r \sum_{s=1}^{\ell-1} f(R_{\ell-s-1}, \Omega) T_s(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) + f(R_{\ell-1}, \Omega) T'_0(\Omega) \mathbf{e}(\Omega) \mathbf{e}^H(\Omega) \quad (91)$$

Recalling that $\mathbf{T}_0(\Omega) = \mathbf{I}_r$ and defining $T'_0(\Omega) = \frac{1}{r}$, we obtain from (91) the scalar $T_\ell(\Omega)$:

$$T_\ell(\Omega) = r \left(\sum_{s=1}^{\ell-1} f(R_{\ell-s-1}, \Omega) T_s(\Omega) + f(R_{\ell-1}, \Omega) T'_0(\Omega) \right). \quad (92)$$

The following equations summarize the recursion in terms of only scalar functions.

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} g(T_{\ell-s-1}, \lambda) R_s(\lambda)$$

$$T_\ell(\Omega) = r \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}, \Omega) T_s(\Omega)$$

$$f(R_s, \Omega) = \frac{\beta}{T_c^2} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 \int \lambda R_s(\lambda) d F_{|\mathbf{A}|^2}(\lambda) \quad |x| \leq \pi$$

$$g(T_s, \lambda) = \frac{r^2 \lambda}{2\pi T_c^2} \int_{-\pi}^{\pi} \left| \Phi \left(\frac{\Omega}{T_c} \right) \right|^2 T_s(\Omega) d\Omega \quad s = 0, 1, \dots$$

with $T_0(\Omega) = \frac{T_c}{r}$ and $R_0(\lambda) = 1$. Let us observe that the different expressions of $g(T_s, \lambda)$ for $s = 0, 1, \dots$ could be absorbed in a unified expression by initialize the recursion with $T_0(\Omega) = \frac{T_c}{r}$ instead of using $T'_0(\Omega) = \frac{1}{r}$.

The recursion in the statement of Theorem 2 is obtained by defining

$$f(R_s) = \int \lambda R_s(\lambda) dF_{|\mathbf{A}|^2}(\lambda)$$

and

$$\nu(T_s) = \frac{r^2}{2\pi T_c} \int_{-\pi/T_c}^{\pi/T_c} |\Phi(\omega)|^2 T_s(\omega) d\omega$$

and by expressing $R_\ell(\lambda)$ and $T_\ell(\omega)$ as recursive functions of $f(R_s)$ and $\nu(T_s)$.

REFERENCES

- [1] L. Cottatellucci, R. R. Müller, and M. Debbah, “Asynchronous CDMA systems with random spreading—part i: Fundamental limits,” *Submitted to IEEE Transactions on Information Theory*, Feb. 2007.
- [2] L. Cottatellucci and R. R. Müller, “A systematic approach to multistage detectors in multipath fading channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3146–3158, Sept. 2005.
- [3] S. Verdú, *Multuser Detection*. New York: Cambridge University Press, 1998.
- [4] J. S. Goldstein and I. S. Reed, “Reduced rank adaptive filtering,” *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 492–496, Feb. 1997.
- [5] J. S. Goldstein, I. S. Reed, and L. L. Scharf, “A multistage representation of the Wiener filter based on orthogonal projections,” *IEEE Transactions on Information Theory*, vol. 44, no. 7, Nov. 1998.
- [6] S. Moshavi, E. G. Kanterakis, and D. L. Schilling, “Multistage linear receivers for DS–CDMA systems,” *International Journal of Wireless Information Networks*, vol. 3, no. 1, pp. 1–17, Jan. 1996.
- [7] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore and London: The Johns Hopkins University Press, 1996.
- [8] D. Divsalar and M. K. Simon, “Improved CDMA performance using parallel interference cancellation,” in *Proc. of Military Communications Conference*, Oct. 1994, pp. 911–917.
- [9] D. Divsalar, M. K. Simon, and D. Raphaeli, “Improved parallel interference cancellation for CDMA,” *IEEE Transactions on Communications*, vol. 46, no. 2, pp. 258–268, Feb. 1998.
- [10] L. G. F. Trichard, J. S. Evans, and I. B. Collings, “Large system analysis of linear multistage parallel interference cancellation,” *IEEE Transactions on Communications*, vol. 50, no. 11, pp. 1778–1786, Nov. 2002.
- [11] R. R. Müller and S. Verdú, “Spectral efficiency of low–complexity multiuser detection,” in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Sorrento, Italy, June 2000, p. 439.
- [12] —, “Design and analysis of low–complexity interference mitigation on vector channels,” *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1429–1441, Aug. 2001.
- [13] W. Hachem, “Simple polynomial detectors for CDMA downlink transmissions on frequency-selective channels,” *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 164–172, Jan. 2004.
- [14] L. Li, A. Tulino, and S. Verdú, “Design of reduced-rank MMSE multiuser detectors using random matrix methods,” *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 986 – 1008, June 2004.

- [15] Laura Cottatellucci, Merouane Debbah, and R. R. Müller, “Asymptotic analysis of linear detectors for asynchronous CDMA systems,” in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Chicago, Illinois, June/July 2004.
- [16] L. Cottatellucci, R. R. Müller, and M. Debbah, “Efficient implementation of multiuser detectors for asynchronous CDMA,” in *Proc. 42nd Allerton Conf. on Communication, Control and Computing*, Monticello, Illinois, Sept./Oct. 2004, pp. 357–366.
- [17] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*. Cambridge, U.K.: Cambridge University Press, 2003.
- [18] L. Cottatellucci, “Low complexity multiuser detectors with random spreading,” Ph.D. dissertation, TU Wien, Vienna, Austria, Mar. 2006.
- [19] D. R. Brown, M. Motani, V. V. Veravalli, V. Poor, and C. R. Johnson, “On the performance of linear parallel interference cancellation,” *IEEE Transactions on Information Theory*, pp. 1957–1970, July 2001.
- [20] D. Guo, L. K. Rasmussen, and T. J. Lim, “Linear parallel interference cancellation in long-code CDMA multiuser detection,” *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 12, pp. 2074–2081, Dec. 1999.
- [21] S. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, ser. Prentice Hall Signal Processing Series. Prentice Hall, 1993, vol. 1.
- [22] D. Tse and S. Hanly, “Linear multiuser receivers: Effective interference, effective bandwidth and user capacity,” *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.
- [23] M. Honig and W. Xiao, “Performance of reduced-rank linear interference suppression,” *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1928–1946, July 2001.
- [24] A. Mantravadi and V. V. Veeravalli, “MMSE detection in asynchronous CDMA systems: An equivalence result,” *IEEE Transactions on Information Theory*, vol. 48, no. 12, pp. 3128–3137, Dec. 2002.
- [25] P. Schramm and R. R. Müller, “Spectral efficiency of CDMA systems with linear MMSE interference suppression,” *IEEE Transactions on Communications*, vol. 47, no. 5, pp. 722–731, May 1999.
- [26] J. Huber, *Trelliscodierung*. Berlin, Germany: Springer-Verlag, 1992.
- [27] P. Billingsley, *Probability and Measure*, 3rd ed., ser. Wiley in probability and mathematical statistics. John Wiley & Sons, 1995.
- [28] R. Speicher, “Freie Wahrscheinlichkeitstheorie,” Lecture Notes, Heidelberg, Germany, 1997/98.
- [29] Y. Q. Yin, “Limiting spectral distribution for a class of random matrices,” *Journal of Multivariate Analysis*, vol. 20, pp. 50–68, 1986.
- [30] K. W. Wachter, “The strong limits of random matrix spectra for sample matrices of independent elements,” *The Annals of Probability*, vol. 6, no. 1, pp. 1–18, 1978.